DOCUMENT RESUME

ED 037 476	24	TE 500 596

AUTHOR	Jewell, Ross M.; And Others
TITLE	The Effectiveness of College-Level Instruction in Freshman Composition. Final Report.
INSTITUTION	Northern Iowa Univ., Ceder Falls.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-5-0803
PUB DATE	69
CONTRACT	OEC-SAE-OE-4-10-053
NOTE	273p.
EDRS PRICE	EDRS Price MF-$1.25 HC-$13.75
DESCRIPTORS	Academic Performance, *College Freshmen, Composition (Literary), *Composition Skills (Literary), Computer Assisted Instruction, English, English Education, *English Instruction, Experiments, Performance Tests, *Research, Sex (Characteristics), *Standardized Tests, Statistical Data, Tables (Data), Verbal Tests

ABSTRACT

This final report of a two-stage project describes an effort to determine whether students receiving instruction in freshman English composition perform better on standardized tests than students who do not receive similar instruction, when both groups are in college the same length of time. The second phase of the experiment detailed in the report involves 1,040 matched pairs of students from the University of Northern Iowa, the University of Iowa, Kent State University, the University of Colorado, and Northern Illinois University. Using the Cooperative English Tests: English Expression (COOP), the College Entrance Examination Board (CEEB) English Composition Test, and a theme as test instruments, the authors include computer-generated test results covering frequent intervals over the 2-year period on: (1) overall performance, (2) performance by ability quarters by sex, and (3) performance by sex. Background, procedures, summary, and recommendations are included. Forty-eight statistical tables reveal performance data. (RL)

TEO

U.S. DEPARTMENT OF HEALTH, EDUCATION
& WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

BR 5-0803
PA 24
OE/BR

ED037476

FINAL REPORT
Project No. 2188, Amended
Contract No. SAE-OE-4-10-053 Amended

THE EFFECTIVENESS OF COLLEGE-LEVEL INSTRUCTION IN FRESHMAN COMPOSITION

Ross M. Jewell, Director
John Cowley, Gordon Rhum

University of Northern Iowa
Cedar Falls, Iowa 50613

with Computer Programming by
William Snider, University of Iowa
Carl Schnittjer, University of Northern Iowa

1969

U.S.DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

TE 500 596

ED037476

FINAL REPORT

Project No. 2188, Amended
Contract No. SAE-OE-4-10-053 Amended

THE EFFECTIVENESS OF COLLEGE-LEVEL INSTRUCTION IN FRESHMAN COMPOSITION

Ross M. Jewell, Director
John Cowley, Gordon Rhum

with Computer Programming by
William Snider, University of Iowa
Carl Schnittjer, University of Northern Iowa

University of Northern Iowa
Cedar Falls, Iowa 50613

1969

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

CONTENTS

Page

Contents 3
List of Tables 7
List of Figures 13

Foreword 15
Summary 17

Background 19
- Statement of the Problem 19
- Selection of Cooperating Universities 20
- Composition Programs in Participating Universities 21
- Related Research 25

Procedures 38
- Establishing Matched Pairs 38
- Evaluative Instruments 40
  - Objective tests 40
  - Theme 40
- Reliability 42
  - Cooperative English Tests: English Expression 42
  - The College Entrance Examination Board English Composition Test 42
  - Theme 43

Findings and Analysis 45
- General Data 45
  - The Samples 45
  - All Tests--September 1964 through May 1966 50
  - Intercorrelation Data 51
- First Semester Sample 54
  - COOP, September 1964-January 1965 54
    - Overall performance 54
    - Performance by ability quarters 58
    - Performance by ability quarters by sex 60
  - CEEB, September 1964-January 1965 62
    - Overall performance 62
    - Performance by ability quarters 64
    - Performance by ability quarters by sex 66
  - Theme Rating, September 1964 and January 1965 68
    - Overall performance 68
    - Performance by ability quarters 71
    - Performance by ability quarters by sex 71

Page

- First Year Sample 75
  - COOP, First Semester 75
    - Overall performance 75
    - Performance by ability quarters 78
  - CEEB, First Semester 80
    - Overall performance 80
    - Performance by ability quarters 80
  - Theme Rating, First Semester 83
    - Overall performance 83
    - Performance by ability quarters 85
  - COOP, Second Semester 85
    - Overall performance 85
    - Performance by ability quarters 88
  - CEEB, Second Semester 90
    - Overall performance 90
    - Performance by ability quarters 92
  - Theme Rating, Second Semester 94
    - Overall performance 94
    - Performance by ability quarters 94
  - COOP, First Year 97
    - Overall performance 97
    - Performance by ability quarters by sex 97
  - CEEB, First Year 100
    - Overall performance 100
    - Performance by ability quarters by sex 100
  - Theme Rating, First Year 103
    - Overall performance 103
    - Performance by ability quarters by sex 106
- Two Year Sample 106
  - COOP Performance, First Semester 106
    - Overall performance 106
    - Performance by sex 106
  - COOP Performance, Second Semester 109
    - Overall performance 109
    - Performance by sex 109
  - COOP Performance, First Year 109
    - Overall performance 109
    - Performance by sex 109
  - COOP Performance, Second Year 112
    - Overall performance 112
    - Performance by sex 112
  - COOP Performance, Two Academic Years 114
    - Overall performance 114
    - Performance by sex 114
  - Summary 116

Page

CEEB Performance, First Semester 116
  Overall performance 116
  Performance by sex 118
CEEB Performance, Second Semester 118
  Overall performance 118
  Performance by sex 118
CEEB Performance, First Year 120
  Overall performance 120
  Performance by sex 120
CEEB Performance, Second Year 120
  Overall performance 120
  Performance by sex 123
CEEB Performance, Two Academic Years 123
  Overall performance 123
  Performance by sex 123
Summary 125
Theme Performance, First Semester 125
  Overall performance 127
  Performance by sex 127
Theme Performance, Second Semester 127
  Overall performance 127
  Performance by sex 127
Theme Performance, First Year 129
  Overall performance 129
  Performance by sex 129
Theme Performance, Second Year 129
  Overall performance 129
  Performance by sex 129
Theme Performance, Two Academic Years 132
  Overall performance 132
  Performance by sex 132
Summary 133

Summary, Conclusions, and Observations 134
  Overall Findings 134
  Findings by Sex 136
    Superiority of female performance 137
    Gains by sex 141
    Theme performance 142
    Performance by ability quarters, by sex 143
  Number of Students 144
  Matched Pairs 145
  Themes as Tests 146

Recommendations 148

References 151

Page

Appendixes

Appendix A - Theme Topics and Instructions 155
Appendix B - Choice of Experimental Design 167
Appendix C - Procedure for Evaluating Themes 173

LIST OF TABLES

Table | Page

I Achievement as of September 1964 of the Sample of 1964-65 Freshman Students and of Various Persisting Portions of that Sample: Combined Institutions 46

II Performance of 122 Matched Pairs of Students on each of Three Criterion Measures at Beginning, Middle, and End of First Year and End of Second Year of College: Combined Institutions 51

III Performance of Available Samples of Matched Pairs of Students on Three Criterion Measures at Beginning, Middle, and End of First Year and End of Second Year of College: Combined Institutions 52

IV Intercorrelations Among Eight Variables for 597 Experimental Students in September 1964 and January 1965: Combined Institutions 53

V Intercorrelations Among Eight Variables for 597 Control Students in September 1964 and January 1965: Combined Institutions 55

VI The Performance of 597 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 56

VII The Performance of 597 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965, by Z-Score[1] Quarters: Means, Standard Deviations, and t-Ratios: Combined Institutions 59

VIII Performance on Cooperative English Tests: English Expression by Matched Pairs by Sex in September 1964 and January 1965 at Each of Four Z-score[1] Levels: Combined Institutions 61

Table Page

IX The Performance of 597 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 63

X The Performance of 597 Matched Pairs of Students on the College Entrance Examination Board English Composition Test in September 1964 and January 1965, by Quarters: Means, Standard Deviations, and t-Ratios: Combined Institutions 65

XI Performance on College Entrance Examination Board English Composition Test by Matched Pairs by Sex in September 1964 and January 1965 at Each of Four Z-score[1] Levels: Combined Institutions 67

XII The Performance of 597 Matched Pairs of Students on Total Theme Rating in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 69

XIII The Performance of 597 Matched Pairs of Students on Total Theme in September 1964 and January 1965 at Each of Four Z-score[1] Levels; Means, Standard Deviations, Differences in Means, and t-Ratios: Combined Institutions 72

XIV Performance on Theme by Matched Pairs by Sex in September 1964 and January 1965 at Each of Four Z-score[1] Levels: Combined Institutions 73

XV The 596 Matched Pairs Who Persisted Through January 1965: Means on September Theme and September Z-score[1] by Sex by Subgroup by Ability Level Established on September Theme Rating 76

XVI The Performance of 365 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 77

Table Page

XVII The Performance of 365 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965, by Z-score[1] Quarters; Means, Standard Deviations, and t-Ratios: Combined Institutions 79

XVIII The Performance of 365 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 81

XIX The Performance of 365 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965, by Z-score[1] Quarters: Means, Standard Deviations, and t-Ratios: Combined Institutions 82

XX The Performance of 365 Matched Pairs of Students on Total Theme Rating in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 84

XXI The Performance of 365 Matched Pairs of Students on Total Theme in September 1964 and January 1965 at Each of Four Z-score[1] Levels; Means, Standard Deviations, Differences in Means, and t-Ratios: Combined Institutions 86

XXII The Performance of 365 Matched Pairs of Students on the Cooperative English Tests: English Expression in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 87

XXIII The Performance of 365 Matched Pairs of Students on the Cooperative English Tests: English Expression in January 1965 and May 1965, by Z-score[1] Quarters: Means, Standard Deviations, and t-Ratios: Combined Institutions 89

Table Page

XXIV The Performance of 365 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 91

XXV The Performance of 365 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in January 1965 and May 1965, by Z-score[1] Quarters: Means, Standard Deviations, and t-Ratios: Combined Institutions 93

XXVI The Performance of 365 Matched Pairs of Students on Total Theme Rating in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 95

XXVII The Performance of 365 Matched Pairs of Students on Total Theme in January 1965 and May 1965 at Each of Four Z-score[1] Levels; Means, Standard Deviations, Differences in Means, and t-Ratios: Combined Institutions 96

XXVIII The Performance of 365 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 98

XXIX Performance on Cooperative English Tests: English Experession by Matched Pairs by Sex in September 1964 and May 1965 at Each of Four Z-score[1] Levels: Combined Institutions 99

XXX The Performance of 365 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 101

Table Page

XXXI Performance on College Entrance Examination Board English Composition Test by Matched Pairs by Sex in September 1964 and May 1965 at Each of Four Z-score[1] Levels: Combined Institutions 102

XXXII The Performance of 365 Matched Pairs of Students on Total Theme Rating in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 104

XXXIII Performance on Theme Total by Matched Pairs by Sex in September 1964 and May 1965 at Each of Four Z-score[1] Levels: Combined Institutions 107

XXXIV The Performance of 122 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 108

XXXV The Performance of 122 Matched Pairs of Students on the Cooperative English Tests: English Expression in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 110

XXXVI The Performance of 122 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 111

XXXVII The Performance of 122 Matched Pairs of Students on the Cooperative English Tests: English Expression in May 1965 and May 1966; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 113

XXXVIII The Performance of 122 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and May 1966; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 115

Table Page

XXXIX The Performance of 122 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 117

XL The Performance of 122 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 119

XLI The Performance of 122 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 121

XLII The Performance of 122 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in May 1965 and May 1966; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 122

XLIII The Performance of 122 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and May 1966; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 124

XLIV The Performance of 122 Matched Pairs of Students on Total Theme Rating in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 126

XLV The Performance of 122 Matched Pairs of Students on Total Theme Rating in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions 128

Table		Page
XLVI	The Performance of 122 Matched Pairs of Students on Total Theme Rating in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions	130
XLVII	The Performance of 122 Matched Pairs of Students on Total Theme Rating in May 1965 and May 1966; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions	131
XLVIII	The Performance of 122 Matched Pairs of Students on Total Theme Rating in September 1964 and May 1966; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: Combined Institutions	133

LIST OF FIGURES

Figure		
1	General Information on 1964-65 Enrollment and Freshman Composition Classes for each Institution	22
2	Composition Programs at Participating Institutions, 1964-65	23

FOREWORD

This is the final report of Research Project 2188, amended, under a contract between the United States Office of Education and the University of Northern Iowa. An interim report, covering a pilot phase of the study, was completed in December 1966.[1] Project 3177, extending the present research for some of the same students through the senior year, is scheduled for completion in 1970.

The investigators express their appreciation to the many individuals and organizations which have assisted in bringing this project to fruition. First among these must be Dr. J. W. Maucker, President of the University of Northern Iowa; others are Vice President William C. Lang; Dr. H. W. Reninger, formerly Head of the Department of English Language and Literature; Dr. Keith McKean, present Head of the Department; and Dr. Marshall Beard, Registrar. Had the administration of the university not had the courage to allow students to omit a course frequently considered to be vital to their success in college and in life, this project could not even have begun. Acknowledgment is also extended to the officers and staff of the cooperating universities, especially Dr. Harold Kelling at the University of Colorado, Dr. Richard Braddock at the State University of Iowa, Dr. Orville Baker and Mr. Gordon Magnuson at Northern Illinois University, and Dr. Kenneth Pringle and Mr. George Whitesel at Kent State University.

For a careful reading of the manuscript, which revealed a number of miscues and oversights so that they could be corrected prior to publication, we are indebted to Dean Emeritus Martin J. Nelson. He should not be held responsible for any which remain.

To Mrs. Alice Prigge, our typist, we are uniquely indebted for her alertness, diligence, and resourcefulness. She aided us in overcoming many difficulties and pointed out shortcomings which might otherwise have persisted through publication.

Finally, the investigators owe a debt of gratitude to the students who participated in the investigation. They were cooperative and helpful, whether they were in the experimental subgroup or the control subgroup. We hope that the impact of the findings reported here will justify the students' cooperation.

---

[1] Jewell, Ross, John Cowley, and Gordon Rhum, Interim Report: The Effectiveness of College-Level Instruction in Freshman Composition (Cooperative Research Project 2188), Cedar Falls, Iowa: State College of Iowa, 1966. The State College of Iowa became the University of Northern Iowa in 1967.

SUMMARY

This is the Final Report of Research Project 2188 as amended. The hypothesis involved was that on tests related to writing, performance of students receiving instruction in freshman English composition does not differ significantly from the performance of similar students not receiving instruction in freshman English composition, when both groups have been in college the same length of time. The research was conducted in two stages. The first stage, at the University of Northern Iowa, began with the fall semester, 1963, and concluded with the spring semester, 1965. The first phase was reported in an Interim Report. The second stage, in which the University of Northern Iowa was joined by the University of Iowa, Kent State University, the University of Colorado, and Northern Illinois University, began in September 1964 and ended in May 1966. The present report concerns the second stage.

In the fall of 1964 the basic pool of 4,190 freshman students from five institutions combined were subdivided randomly, within sex and ACT English score, into experimentals (N=1,408) and controls (N=2,782). The experimental students did not enroll in freshman composition courses; the control students did. Following testing at the beginning of the fall semester of 1964, 1,040 matched pairs of students were formed on the basis of sex, age, and scores on the Cooperative English Tests: English Expression (COOP), College Entrance Examination Board English Composition Test (CEEB), and a theme. The three tests were again administered at the end of the first semester, second semester, and fourth semester. Numbers of fully described matched pairs who persisted were, respectively, 597, 365, and 122.

Results partially confirmed and partially denied the hypothesis. Of nine main comparisons--COOP, CEEB, and theme at the end of the first semester, the end of the second semester, and the end of the fourth semester--the null hypothesis was denied on three, the control students performing significantly better than the experimentals on COOP and theme at the end of the first semester, and on COOP at the end of the second semester.

Test scores were also analyzed in terms of sex and ability level of students. Females performed consistently better than males, particularly at the lowest one-quarter of ability. Teachers and researchers should not overlook this differential between the sexes in performance on tests related to composition. It was also at the lowest ability level that there were the strongest indications of some superiority of the control subgroup over the experimental subgroup.

Data were also analyzed for each of the three criterion measures for the 122 pairs who were available at the final testing date. In this analysis--using a constant N of 122 instead of diminishing N's (597, 365, 122)--none of the nine comparisons between experimental subgroup and control subgroup yielded a significant difference in means.

BACKGROUND

Statement of the Problem

Research in college composition has not been plentiful, and most of the studies reported have concentrated on comparing some innovation with a standard procedure. Variables have ranged from the number of student papers written through the amount of teacher comment on each paper to the influence of such subjects as rhetoric and grammar on the performance of the student. In every case the other element in the comparison was the particular arrangement of freshman composition at the institution in which the research was done. Seldom has a statistically significant difference appeared, and the difficulty is that, even where it has, the difference has been between a particular innovation and what might be termed standard procedure. A tacit assumption in such research has been that the "standard" course improved student writing and the question was whether the innovation would produce a result different from that produced by the standard course. These investigations seldom included comparisons of the results with an arrangement involving no formal instruction in English composition.

A second difficulty with the research reported has been that the statistical comparisons involved a relatively small number of students. The question is always present as to whether the sample employed is sufficiently large and broadly based to be reasonably representative of a given group--for example, all entering college freshmen in a substantial number of American colleges. In those few instances in which a statistically significant difference has been found, the degree to which generalizations beyond the samples investigated may be made is uncertain.

The present investigators decided to attempt to overcome both of these deficiencies. They planned to compare students who had received no instruction of the sort generally given in freshman composition with comparable students who had received such instruction. In order to develop statistics for a reasonably broad and a reasonably diverse population, they planned to engage several institutions in replicating the experiment. This procedure would give a numerical, geographical, and academic variety to the population. If the results in the participating institutions were in substantial agreement, the conclusions could be stated with considerable force.

The goals of the investigation, then, were to test two hypotheses:

(1) That the writing performance of the students enrolled in a freshman composition sequence is not significantly

different from the writing performance of students not enrolled in a freshman composition sequence when the two groups have been in college for an equal length of time.

(2) That the results obtained in (1) will be present in many colleges or universities.

A by-product of the testing of the hypotheses would be the accumulation of statistics based upon a reasonably large and diverse sample of students who had received no instruction in college freshman composition. Such a set of statistics might prove useful in providing a realistic and stable base for investigating the effect of innovation as well as of the "standard" course itself. Meaningful use of these statistics could be made only if the investigators testing an innovation utilized the evaluative instruments employed in the present investigation.

The investigation was divided into a pilot phase, conducted at the University of Northern Iowa* from 1963 to 1965, and a major phase which ran from September 1964 through May 1966. The major phase is reported in this document. The results of the pilot study are available in the interim report. Procedures followed in the pilot phase were replicated at the University of Northern Iowa and at four other universities: the University of Colorado, the University of Iowa, Kent State University, and Northern Illinois University. Each of these institutions has been assigned, randomly, a number from one to five. Future references will be by these numbers and not by the names of the institutions.

Selection of Cooperating Universities

The United States Office of Education authorized a total of six institutions in the experiment. The investigators originally intended to include institutions that were varied in size and type: large, small, liberal arts, engineering, private, public, and so forth. Since a freshman class of close to 1,100 would be necessary to ensure sufficient retention, after two years, the total enrollment of each institution had to exceed 4,000. Some effort was made to include geographical distribution also. Time became a factor in the selection because the project was approved in April 1963, and the schools had to be selected in the fall of 1963, in order to enable them and the investigators to plan adequately for the start of the project in the fall of 1964.

---

*On July 1, 1967, the State College of Iowa became the University of Northern Iowa. Thus the whole investigation was completed while the school was called a college, while the report is being written under the new name. University of Northern Iowa will be used throughout.

Of the thirty schools that responded to the first letter, only five remained after the second round of correspondence, all public universities: the University of Colorado, the University of Iowa, Northern Illinois University, and Kent State University.*

Composition Programs in Participating Universities

Figures 1 and 2 present data concerning these participating institutions and their freshman composition programs. Institutions 2-5 were of similar size in 1964-65, all being more than twice the size of University 1. Total freshman composition enrollment, fall, 1964, ranged from about 1,100 to about 4,111. Men constituted almost exactly one-half of the freshman enrollment at four institutions, but only 38 percent at the fifth. Graduate assistants were used for instruction at 3 of the 5 institutions. Teaching loads for full-time staff varied from 9 to 12 hours and class size from 22 to 30 students. Some form of exemption from composition, and some method of optional, outside-of-class help were available at all institutions. Across-the-board class sectioning was the practice at one institution, sectioning for only high students at two, and no sectioning at two. Credit allowances for composition varied from 5 to 8 semester hours.

There was variety in the content of the programs. All emphasized exposition in the first semester or quarter, one emphasizing it for the year. Other common emphases the first semester were organization, central idea, and sentence structure. One institution included argument the first semester. Variety was greater in the second semester or quarter than in the first: the one institution continued exposition, two stressed argument, one imaginative writing, and one literary analysis; three included some literature during the year, two did not. Three institutions required research papers in the second semester or third quarter, two did not.

There were differences in matters other than content. The number of themes for the year varied from 16 to 22. The number of in-class themes varied from 2 to 8, with one university not reporting that item. Data concerning theme length were incomplete. Average theme lengths reported for the first semester were 300 to 500 words, for the second semester 400 to 950 (the latter including a 2,000-word research paper), and for the third term 1,250, including a research paper.

---

*The sudden death of Dr. Herbert Hackett of the State University of New York at Buffalo after the project was underway led to the elimination of that institution.

	Institution				
	One	Two	Three	Four	Five
Total Enrollment	5,519	13,380	13,252	12,672	14,480
Total Freshman Enrollment Sept. 1964	1,914	2,800	4,563	4,842	3,171
Total Freshman Enrollment May 1965	1,649	2,700	4,400	3,849	2,974
Approximate Percent of Men Freshmen, September 1964	38	50	50	50	53
Number of Instructors in Freshman Composition Full Time	14+	20	48+	41	12
Number of Graduate Assistants as Instructors in Freshman Composition	0	23	8	0	84
Normal Teaching Load	9 hrs.	10-12 hrs.	10-12 hrs.	9 hrs.	10-12 hrs.
Average Class Size	30	25-26	25-30	27	22
Exemption from Freshman Composition Possible?	Yes	Yes	Yes	Yes	Yes
Classes Sectioned by Ability?	High only	No	No	High only	Yes*
Optional Outside-of-Class Help Available?	Yes	Yes	Yes	Yes	Yes

+Two part-time instructors also.
*Sections determined by Placement Test Scores.

FIGURE 1

GENERAL INFORMATION ON 1964-65 ENROLLMENT AND FRESHMAN COMPOSITION CLASSES FOR EACH INSTITUTION

	Institution				
	One	Two	Three	Four	Five
Length of Program	2 sem.	2 sem.	3 qr.	2 sem.	2 sem.
Credits	5 sem.	6 sem.	9 qr.	8 sem.	8 sem.
Emphasis - Term I	Exposition, Development, Sentence Structure, Conventions, Lang. Study, Organization.	Reading & Writing Expository Prose, Organization, Sentence Structure, Central Idea.	Exposition, Opinion, Argument, Organization, Clarity, Precision.	Exposition,[a] Central Idea.	Speaking, Reading, Writing, Listening, Written & Oral Exposition.
Emphasis - Term II	Exposition, Effectiveness & Style, Semantics.	Argument, Logic, Rhetoric.	Imaginative & Emotional Writing.	Literary Analysis,[b] Research Paper.	Reasoning, Argument, Criticism, Research.
Emphasis - Term III			Research, Literary Analysis.		
No. of Themes Term I	10	10	6	10	11[c]
No. of Themes Term II	8	8[d]	6	7	11[c]
No. of Themes Term III			4		
No. of In-class Themes Term I	2	0	2/3	5	
No. of In-class Themes Term II	1	2	2/3	3	

[a] Final grade determined by 500-word theme and objective test.
[b] Final grade based 2/3 on writing, 1/3 on literary analysis.
[c] Eight speeches in addition.
[d] One must be revised.

FIGURE 2

COMPOSITION PROGRAMS AT PARTICIPATING INSTITUTIONS, 1964-65

No. of In-class Themes Term III	0	0	0	0	0
Theme Length Term I	300 words		500 words		
Theme Length Term II	400 words	6-8 pp.	667 words	950 words[e]	
Theme Length Term III			1,250 words		
Research Paper Required	No	No	Yes	Yes	Yes
Literature Term I	No	No	No		No
Literature Term II	No	No	Yes	Yes	No
Literature Term III			Yes		

[e] Including a 2,000-word research paper.

FIGURE 2
CONTINUED

Special mention should be made of the program at one institution--a communications approach, combining reading, writing, speaking, and listening. Eleven themes and eight speeches were required each semester. There was emphasis on exposition and argument, a research paper was required but no study of literature was included. This program differed materially from all the others.

Therefore, several major types of composition programs were included in this study: communications approach, stress on exposition only, stress on exposition and argument, and stress on exposition and literary analysis, some with and some without research papers. The programs involved are representative of those at many state universities requiring composition.

Related Research

No research has come to the investigators' attention which is directly comparable to the present study. Nearly all the research compares some innovation with a standard procedure. Such studies ordinarily vary the frequency of writing in the composition course as the experimental variable. Most of these obtained no statistically significant differences in the performance of the groups of students at the end of instruction. A summary of projects with some relevance to the current study is given below.

Arnold, Lois. *Effects of Frequency of Writing and Intensity of Evaluation upon Performance in Written Composition of Tenth Grade Students* (Cooperative Research Project Number 1523), Tallahassee: Florida State University, 1963, University Microfilms No. 63-6344.

Miss Arnold conducted her research in 1961-1962 at two Florida high schools, in each of which a teacher was scheduled to teach four groups of students in the tenth grade. The four groups at each school were average classes, determined by sectioning on the basis of scores on the following tests: *Pintner General Ability Test*, *Metropolitan Achievement Battery*, *School and College Ability Test*, and *Differential Aptitude Tests*. Students were classified as low average, middle average, or high average on the basis of the *DAT* scores. Nothing is said of student-to-student matching. The experiment lasted for the school year. Each teacher at each school used four teaching methods, a different one for each of her four classes as follows:

1. Infrequent writing, moderate evaluation: one theme, approximately 250 words, each six weeks. Evaluation was concentrated on one matter each time: once on sentence structure, once on organization, etc.

2. Frequent writing, moderate evaluation: some writing four times a week, varying from two sentences to two pages or more. The evaluation was handled as in 1 above.

3. Infrequent writing, intensive evaluation: one theme each six weeks, approximately 250 words. Every error in usage, sentence structure, and mechanics was marked and detailed comments written on the paper. Students corrected all errors, revised or rewrote until the paper was satisfactory.

4. Frequent writing, intensive evaluation: one 250-word theme weekly, evaluated meticulously as in 3 above (pp. 40-2).

Two evaluative instruments were used, _STEP Essay Tests_ and _STEP Writing Tests_, the former a writing test, the latter an objective test. Both were administered at the beginning and at the end. Three experienced (former) English teachers independently rated the _STEP Essay Tests_, the pretests in December and January, and the post-tests in May and June.

Miss Arnold reached four conclusions:

1. There is no assurance that intensive evaluation is any more effective than moderate evaluation in improving the quality of written composition.

2. It must not be assumed that frequent practice is in itself a means of improving writing.

3. There is no evidence that any one combination of frequency of writing and intensity of evaluation is more effective than another.

4. There is no indication that frequent writing and intensive evaluation are any more effective for one ability level than are infrequent writing and moderate evaluation (p. 62).

In this study there was no significant difference between the performances of men and women.

The University of Northern Iowa investigators wonder whether graders might have evaluated more alike had they conferred on an occasional paper (four correlations were in the 0.50's, the others being 0.62 and 0.76), and why, in a gains study, all themes were not scored at a single time with prethemes and post-themes mixed. A table showing comparisons of the terminal data only would also have been helpful. That is, how did the groups compare at the end?

Buxton, Earl W. "An Experiment to Test the Effects of Writing Frequency and Guided Practice upon Student's Skill in Written Expression." Unpublished doctoral dissertation, Stanford University, 1958. University Microfilms 58-3596. [As reported in Braddock, *et al.* *Research in Written Composition*. Champaign, Illinois: NCTE, 1963, pp. 58-70.]

This experiment involved 257 students in the University of Alberta who were enrolled in a special "one-year 'emergency' course designed to train teachers for Alberta schools." All 257, who constituted the entire enrollment in the emergency program, carried the same courses (a "canned" schedule). The total group was divided into six classes: two *control* classes, in which students did no extra, out-of-class writing; two *writing* classes, in which students wrote a 500-word paper each week as an extra out-of-class assignment for a total of sixteen weeks; two *revision* classes, in which students did the same amount of writing on the same assignments as the *writing* classes. *Writing* classes were not required to write on the assigned topic and received only a brief paragraph of teacher comment at the end of each theme; there was no marking of errors nor commenting in the margin, and students were not asked to do anything with the papers after getting them back. The *revision* classes were required to write on the assigned topic and papers were marked in terms of unity, organization, logic, correctness, and such matters, with a general comment at the end. Students in the *revision* classes were asked to correct and revise their papers in class on the day the papers were returned and discussed. The teacher was present to give aid.

Criterion measures were two parts of an earlier edition of the *Cooperative English Tests*: "Mechanics of Expression" and "Effectiveness of Expression" (alternate forms before and after), and a theme. Each of two readers assigned a "content" score and an "error" score to each theme. The content score was based on fifteen factors with some factors weighted more than others. A maximum potential score was allotted for each factor. Each reader determined how much of that maximum to assign to that factor in each paper. The error score was determined by counting errors in spelling, punctuation, or mechanics. The points assigned for each of the fifteen factors in a paper by each reader were added; then the count for errors was subtracted from that. The scores for the two readers were averaged, and that mean was arbitrarily divided by three to get a usable scaled score.

The results of Buxton's study show that the revision students--those whose papers were carefully marked and who were required to revise them--made a significantly greater gain in writing achievement as measured by the themes during the seven months of the study than did the writing students--those who wrote the papers but did not revise them. There was a more significant difference in gain scores between the revision students and the control students, who wrote none of the themes; this difference favored the revision students. Concomitant conclusions: theme ratings are reliable if the raters are thoroughly practiced in their system and frequently check on what they are doing, and (since there was no significant difference between the groups on the objective test scores) the theme ratings in this study measure something that the particular objective test used did not measure.

It is not clear whether the division into groups took into account the balance of men and women. If, for example, the revision classes had more women than either of the other two groups, that could affect the results.

Heys, Frank, Jr. "The Theme-a-Week Assumption: a Report of an Experiment," English Journal, 51 (May 1962), 320-22.

This experiment dealt with varying the amount of writing and the amount of reading in high school English classes. Two classes in each of the four high school grades were "as closely matched as was possible under the normal sectioning practices of the school." The two classes in each grade were taught by the same teacher; one was designated as the writing class and the other as the reading class. Students in each writing class wrote a theme a week. After it was closely graded, the students corrected or rewrote it. Students in each reading class wrote a theme every three weeks, and spent one class day a week reading books of their own choice. Nothing is said concerning grading or rewriting of the reading-class papers. Evaluation instruments consisted of the STEP writing test and a theme, one of each administered at the beginning and at the end of the experiment. The themes were evaluated by three ETS readers using a nine-point scale.

The students in reading classes achieved a slightly greater improvement in writing scores than did those in writing classes. Generalizations arrived at by the investigator:

1. Frequent writing practice probably yields greater dividends in grade 12 than in grades 9, 10, 11.

2. Frequent writing practice probably yields greater dividends with low groups than with middle or high groups.

3. Frequent writing practice with low groups probably yields greater dividends within the area of content and organization than within the area of mechanics or of diction and rhetoric.

4. The claim that "the way to learn to write is to write" is not substantiated by this experiment.

5. The claim that ability to write well is related to the amount of writing done is not substantiated by this experiment.

6. For many students reading is a positive influence on writing ability.

7. The influence of reading on the ability to write appears to be a separate factor, not directly related to the teacher's personality and enthusiasm (p. 322).

It is not clear how the fourth generalization is supported by the experiment. Since all students in the experiment wrote themes, how can it be inferred that the data failed to support the notion that students learn to write by writing? Furthermore, Heys does not indicate whether the improvement mentioned was statistically significant.

Kincaid, Gerald L. "Some Factors Affecting Variations in the Quality of Students' Writing." Unpublished doctoral dissertation (Michigan State University, 1953). University Microfilms No. 5922.

This experiment attempted "to determine whether a single paper written on a given topic at a particular time [italics Kincaid's] can be considered as a representative sample of his [the student's] writing ability--and thus provide a valid basis for evaluating ability at any time in a writing course." It is of interest, not because it deals with a directly related problem, but because it has implications for any study using theme readers to evaluate results. A group of eighty college students was divided into four subgroups, each of which wrote two papers in one two-hour session on the same day and another two papers in a similar session a week later. Three topics were used: Groups A and C wrote on topics 1 and 2 each time (both argumentative); groups B and D wrote on topics 1 and 3 each time (one argumentative, one expository). Groups A and B wrote each time without examination pressure (papers not counted toward grade); groups C and D wrote without pressure once, and with it the other time (papers counted on term grade the first time and not counted on term grade the second time). Papers were rated by three instructors selected from the freshman staff, the rating being made on a ten-point scale (1 unsatisfactory, 10 superior) on each of five categories: grammatical conventions, sentence structure, diction, organization, and content. The score for a paper could lie between 10 and 50; it was determined by computing the mean of the two closest ratings; if the two extreme ratings were equidistant from the middle rating or if the two closest ratings were more than five points apart, the mean of all three was used.

Kincaid drew the following conclusions from this study:

> 1. . . . the findings from this study cast considerable doubt upon the justification of the customary practice of using five letter-grades to designate [individual] achievement in a writing course when a single paper provides the basis for that designation (p. 97).
>
> 2. If an evaluation of overall or average improvement is all that is desired, it can be obtained from a single sample of each student's writing for a pretest and a post-test. . . (p. 99).
>
> 3. . . . in order to develop a program for evaluating individual student improvement in writing (for strong as well as for weak students), it would be advisable to obtain several samples of writing by each student--samples of writing on different topics on the same day and on the same topics on different days. And such samples should be obtained for both the pretest and the post-test (p. 99).

Two matters impress the present investigators: 1) The theme topics used by Kincaid were simpler than those used in the University of Northern Iowa investigation. If more difficult topics had been used by Kincaid the results might have been different. 2) The findings of the Kincaid investigation support the use of group average scores on a single pretheme and a single post-theme.

Kreisman, Arthur, et al. Pilot Study in English. Mimeographed report and dittoed summary of statistics. Ashland, Oregon: Southern Oregon College, 1963 (no pagination).

This is the report of a pilot study designed "to investigate techniques and writing skills as a possible means of establishing the basis for a more extensive research program." It is interesting because the results led the Oregon investigators to abandon further experimentation, and because one of those investigators suggested a study like the University of Northern Iowa study. In the Oregon study, both college freshmen and high school students were involved. Control and experimental groups were matched at both levels: the 89 college students on the Verbal and Quantitative scores on SAT, the total score on SCAT, and the sum of two ratings on the STEP Essay Test; the 108 high school students on the score on the California Test of Mental Maturity and the sum of two ratings on the STEP Essay Test. Both control and experimental students were in each class. The control students wrote a theme a week (a total of 9 for the college group, 36 for the high school group); the experimental students wrote a theme a month (a total of 3 for the college group, 10 for the high school group). Evaluation was based upon comparison of the STEP Essay ratings at the beginning and at the end of the experiment.

There was no significant difference between the college experimental and control groups. The results for the high school groups varied. There was a significant improvement for the below-average high school students in the control group (more writing); there was a slight (non-significant) drop in achievement for the above-average students in the control group (more writing). There was no significant difference in the experimental group (less writing). Dr. Cloer, the statistician, wrote: "It would appear that the principal beneficiaries of the experience in writing were those subjects of below-average ability or those who might be called 'under-achievers,' . . . ."

Comments quoted from Kreisman:

1. No adequate instrument for testing [composition] seems available.

2. The difficulty of obtaining a sufficient number of students to make the experiment valid was one of the major obstacles.

3. . . . a purely quantitative experiment has little chance of being valid.

4. . . . one term of writing practice is not sufficient to form a foundation for judgment regarding the development of writing ability.

5. . . . frequency may indeed be a factor in the development of writing ability.

6. . . . all experiments of this nature are of no value and invalid on an a priori basis.

In the light of the University of Northern Iowa study, the following additional comments are of special interest, the first by Kreisman, the second by Cloer, the statistician: "The emphasis that we thought might be fruitful [for future research] would be one which dealt with student-teacher relationships or with maturation of students regardless of the courses they took," and "Perhaps a better 'experimental group' would be one that did no writing (in English classes) over the experimental period."

McColly, William and Robert Remstad. _Comparative Effectiveness of Composition Skills Learning Activities in the Secondary School_ (Cooperative Research Project 1528). Madison: University of Wisconsin, 1963.

This study attempts to answer three questions:

> Does more writing alone result in better writing?
>
> Do more of "functional non-writing composition learning activities" (practical instruction: working with student-written papers, emphasizing spelling, proof-reading, revision, etc.; group discussion; teacher evaluation and comment) result in better writing?
>
> Does tutoring with immediate feedback (having the teacher present while the writing is being done and advising the student during the process) result in better writing? (p. 18)

To answer the first question, dealing with the effect of the quantity of writing on improvement in writing, the investigators used two classes in the eighth grade and two classes in the ninth grade. To answer the questions relating to "functional non-writing activities" and immediate feedback (tutoring), three classes in each of the tenth, eleventh, and twelfth grades were used. Covariance techniques and, to the extent possible, random selection of samples were employed.

To explore the effect of the amount of writing on improvement in writing, *control* classes in the eighth and ninth grades wrote a theme a month; *experimental* classes wrote a theme a week. All other class activities and assignments were the same. During the year, the eighth-grade *control* classes wrote 9 themes and the eighth-grade *experimentals* wrote 35 themes. The ninth-grade *control* classes wrote 8 themes, the *experimentals*, 34.

To study the effect of non-writing activities and tutoring, one *control* class (a monthly theme with functional instruction), and two *experimental* classes (weekly theme and functional instruction), were organized at each grade level. About 9 writing tasks with functional activities were completed in the *control* classes, about 34 in the *experimental* classes. There were no individual conferences or "tutoring" activities in the first of these *experimental* classes in each grade. There were about 27 regular "tutoring" sessions in the second *experimental* class in each grade. Thus, a ratio of 4-1 was maintained in writing tasks with functional activities between the *experimental* and *control* classes.

Criterion and covariate measures for all students in the experiment included: *SCAT* (IA, IIA, IIIA), Nelson-Denny Reading, *ITED* ("Correctness and Appropriateness of Expression" and "Ability to Interpret Literature"), previous English GPA, overall GPA, and writing samples, two written before the experiment and two written at the end.

Based on this experiment, the answer to the first question is *no*. Results indicated that increase in the amount of writing by itself has no significant effect upon the writing proficiency of high school students. Again, based on this experiment, the answer to the second question is affirmative; the answer to the third question is negative. *Experimental* classes with weekly theme and functional instruction improved significantly compared to the *control* classes. The *experimental* classes with tutoring scored, at the end of the experiment, about half way between the *control* classes and *experimental* classes without tutoring.

Rohman, D. Gordon and Albert Wlecke. *Pre-writing: The Construction and Application of Models for Concept Formation in Writing* (Cooperative Research Project No. 2174), East Lansing, Michigan: Michigan State University, 1964.

This is one of the very few studies that have resulted in a statistically significant difference between *control* and *experimental* groups. Six sections of a college sophomore course in expository writing with an emphasis on pre-writing activities constituted the *experimental* groups. Six sections of a college sophomore course in expository writing with an emphasis on pre-writing activities constituted the *experimental* group. Three sections were taught each quarter for two quarters. The rest of the students enrolled in the same course (11 sections in the Winter term, 10 in the Spring term), constituted the *control* group. The total number of students involved in the experiment is not disclosed. The *experimental* course contained six units: 1. The role of the writer. 2. The escape from category (the concrete rather than the abstract). 3. The escape from cliche (avoiding someone else's way or words). 4. Dynamic relationship to the subject (an urgency to express what the writer has "discovered"). 5. Concrete analogy (expressing one's "discovery" by comparison with something like it). 6. Refinement (finishing the essay). Three major techniques were used: keeping a journal, meditation, and use of analogy. The *control* sections were taught as each teacher wished to teach them, with the exception that all instructors of the *control* sections assigned two 500-word themes on topics used in the *experimental* sections. These themes were used in the evaluation.

Evaluation of the experiment involved four devices: 1. statements written by students in answer to the question: What did you like or dislike about the course?, 2. statements by the teachers who taught the course, 3. "objective" evaluation by readers who did not teach the course, and 4. "subjective" evaluation by teachers who did not teach the course. No objective testing was reported.

Evaluation by students was strongly favorable. Major items were that the course was enjoyed, that it developed freedom in writing and in the discipline of writing and thinking, that criticism of student writing led to involvement in the process of writing, that attitudes toward writing had changed (regarding, for instance, the relationship between thinking and writing), that the use of analogy led to greater concreteness and clarity. Negative criticisms, which were relatively few, included the following: the course was too short; it was too piecemeal; not enough grades were given; class criticism was too negative; the journal was an invasion of privacy; the use of analogy was mechanical.

Instructors gave a number of reactions to the experiment, but their enthusiasm tended to center on three matters: the journal as a device to stimulate students to meditate about their experiences as well as to formulate their meditations in writing, the emphasis on the pre-writing process, and the freshness and soundness of the writing done.

The essays for "objective" evaluation were selected from the total submitted by *control* and *experimental* subgroups on the two topics used by both subgroups. There were 226 *experimental* and 409 *control* essays evaluated. No information is given concerning how these essays were selected. Essays were judged on a four-point scale: 4. superior, 3. above average, 2. below average, 1. incompetent. Three standards, unity, coherence, and emphasis, were guides for the readers. There were eleven readers, four high school teachers and seven college teachers. They worked in teams of eight, three who read at the first session not reading at the second, and three others substituting for them at the second. Each theme was read twice. About 85 percent of the grades assigned were either the same for each theme or only one point different, indicating that the grading was relatively reliable. The results showed a statistically significant difference between the *experimental* and *control* groups in favor of the *experimentals*.

Four members of the English staff not involved in the experiment read the papers "subjectively." They were given a randomly selected sample of 50 *experimental* and 50 *control* themes. Rohman and Wlecke informed these readers concerning which set was *experimental* and which was *control*. Some investigators would not have done that. The readers were asked to answer a series of three questions: "Which set of essays seems to have more originality and in what ways? Generally, in which set of essays does it seem more important for the writers to express themselves and not be misunderstood? Which set of essays gives the greater sense of form?" (pp. 130-1) In addition, the readers were asked a series of specific questions concerning only the *experimental* essays, such as: "Do the techniques employed in the *experimental* essays--the meditation in the 'Loneliness' essays, and the analogy in the 'Coming of Age' essays--seem to provide a more coherent means for the instructor to gauge the success or failure of an essay?" All four readers gave the *experimental* group of essays the higher rating.

Rohman and Wlecke leave so many questions unanswered that the report is difficult to interpret. How many students were in each sample? Were the students of the *experimental* sections similar in ability to those in the *control* sections? Did either sample have appreciably more women than the other? How were the themes that were evaluated selected? Do the 226 *experimental* themes represent a sampling comparable to the 409 *control*? Would a sampling of the *control* students have written as enthusiastically of their course as the *experimentals* did? To what degree did the Hawthorne effect operate? What implications has this study for composition programs generally?

Sutton, Joseph T. and Eliot Allen. The Effect of Practice and Evaluation on Improvement in Written Composition. (Cooperative Research Project No. 1993). Deland, Florida: Stetson University, 1964.

This study randomly divided college freshmen into five groups. The first two of these (Groups I and II) served as controls. During the period of the experiment, these two groups received no instruction in composition and wrote no papers except the six criterion themes which provided the "before" performance and the six criterion themes which provided the "after" performance. Group I wrote all twelve themes within a four-week period at the beginning of the semester. Group II wrote the first six criterion themes the first two weeks of the semester and the second six criterion themes the last two weeks of the semester. Groups III through V were the experimental groups, and all wrote six criterion themes the first two weeks and another six the last two weeks (as did Group II). In the ten-week interval between the writing of criterion themes, Group III wrote no papers but did evaluate four peer papers each week; Group IV wrote one theme each week which was evaluated by the members of Group III; and Group V wrote one class theme each week which was evaluated by a "professor."

Five readers read each theme twice, once to rate it, once to rank it in an order of excellence relative to the other eleven themes by each writer. Rankings were based on five criteria: ideas, mechanics, wording, form, and flavor, each one of which was scored on a five-point scale. A total for the six "before" themes for each student as graded by all five graders, divided by thirty (6 themes x 5 graders) gave an average score for each writer. The same was done for the six "after" themes, and the averages were compared.

Particularly in relation to the University of Northern Iowa study, Sutton and Allen's enterprise is interesting. First, none of the students in any of the groups received direct instruction in composition. Such instruction as Groups IV and V received came from the marks and comments on their papers. Group III gained experience in editing, though uninstructed in the procedure. Groups I and II had no experience whatsoever with composition except the twelve criterion themes. Thus, to a degree this study is similar to the present one in that no direct instruction in freshman composition was given and that some of the groups wrote only the criterion themes. It is different from the present study in that there was not a direct comparison between those completing a freshman program of writing instruction and others not in the freshman English course at all.

The results in the Sutton and Allen study showed an unusual inconsistency between the themes and the objective tests. In theme performance, the members of the five groups showed a significant decline during the experimental period. A decline was observed for the five groups combined and for each group separately. This decline was, of course, unexpected. The authors, in speculating about its source, state: "Unfortunately, it appears that the very procedure necessary to secure such stability [among the theme performances] introduced other factors that may have had a deleterious influence on the results." The frequency of writing of test themes which were neither returned to the student nor commented on seems, in the opinion of Sutton and Allen, to have created an attitude of boredom and impatience among the students. On each of the two objective tests, the Cooperative English Tests: English Expression and the College Entrance Examination Board English Test, the students showed significant improvement. This was true for the five groups combined, and there was no significant variation among the five groups in this respect.

Wolf, Melvin H. Effect of Writing Frequency upon Proficiency in a College Freshman English Course. (Cooperative Research Project 2846), Amherst, Massachusetts: University of Massachusetts, 1966.

This study involved six "regular" sections of college freshman composition and four remedial sections. Two of the regular sections, designated experimental-high frequency, wrote 39 themes in the school year; two sections, designated experimental-low frequency wrote 8 themes in the year; two sections, designated control, wrote 15 themes in the year, the usual number in freshman composition at the University of Massachusetts. Two remedial sections, designated experimental-high frequency, wrote 20 themes in one semester; the other two, designated control, wrote 8 themes in one semester. These themes were carefully evaluated by the instructors and were revised and resubmitted by the students. The objective test used was Cooperative English Tests, Form 1C. Six themes were used as tests: two written at the start, two at the end of the first semester, and two at the end of the second semester. The remedial students, being in the study only one semester, wrote only the first four test themes. Evaluation of the test themes was done by ten instructors under the direction of an experienced instructor who had been a reader for the Educational Testing Service. Wolf drew two conclusions: 1) writing proficiency did not improve with the increase in frequency of writing, 2) there was a high correlation between the scores on objective tests of grammar and mechanics and scores of themes as determined by the reading team. Since COOP has a section on mechanics and a section on effectiveness but usually yields a single score, it is not clear how the second conclusion was arrived at.

PROCEDURES

The overall design of the project involved selecting experimental and control subgroups at each university and testing them on four different occasions: the beginning of the freshman year (September 1964), the end of the first semester (January 1965), the end of the first year (May 1965), and the end of the second year (May 1966). Members of the experimental subgroup received no instruction in freshman composition; members of the control subgroup did receive instruction in freshman composition. The performance of these subgroups was compared at each testing period to determine whether the observed differences in their performance on the criterion measures were statistically significant. Care was taken that the members of each subgroup at each university would be representative of the total freshman class entering that university in September 1964. Members of both experimental and control subgroups pursued a normal academic program except that the experimentals omitted the freshman composition course. The experimental subgroups took other courses instead of freshman composition, usually other general education courses, or courses in the major or minor.

Establishing Matched Pairs

Procedures for establishing matched pairs of students were developed to supply a number of pairs at the start sufficient to assure that after the attrition of two years enough pairs would remain to enable the investigators to draw sound conclusions. These procedures were predicated upon an incoming freshman class of 1,100 students, the approximate size of the freshman class at the smallest of the five universities in September 1964, and large enough to guarantee at least 300 matched pairs at the start. One-third of the 1,100 at each institution were designated experimental students and not permitted to enroll in freshman composition (experimental pool); the remaining two-thirds were designated control students and required to enroll in freshman composition (control pool). The selection of the students from the experimental pool necessarily antedated their actual enrollment in September 1964 in order to assure that they would not be enrolled in freshman composition.

It was necessary, in those institutions which would enroll more than 1,100 freshmen, to devise a procedure which would reduce the potential participants to that number before the selection of the experimental and control pools was made. When approximately 85 percent of the expected, new, beginning freshman students had been cleared for admission at each such institution, 1,100 students were selected from the total group by a random process.

When the group of 1,100 was established at each institution, the next step was to select a subgroup which would include approximately one-third of the 1,100, would contain a ratio between men and women representative of the total group at each university and would reflect the range of performance of that group on the English section of the ACT (at four institutions), or the verbal section of the SAT (at one institution). First, the students were divided by sex, and then within each sex, ranked from high to low in terms of standard scores on the English section of the ACT or the verbal score of the SAT. By use of a table of random numbers, the investigators selected 33 percent of the students of each sex at each score level. The students thus identified at each university became the experimental pool; the remainder of each 1,100 became the control pool.

Matching for all schools was performed at the Data Processing Center at the University of Northern Iowa after the September 1964 testing. After the themes had been scored, matched pairs were formed for each participating institution. Criteria for matching were age, sex, theme score, and a score representing combined performance on the CEEB and the COOP. Students were matched exactly on sex and theme score, within one year on age, and within three points on the combination of CEEB and COOP (Z-score).

The matching may be illustrated from actual data from three pairs of students. The numerals represent, in order, the student's sex (1 for male, 2 for female), total theme score (sum of two ratings), year of birth, and combined objective test score.

Subgroup	Sex	Total Theme Score	Year of Birth	Z-Score
Experimental	2	10	1945	111
Control	2	10	1946	111
Experimental	1	6	1945	85
Control	1	6	1946	85
Experimental	1	11	1946	111
Control	1	11	1946	113

The combining of the scores of the two objective tests was accomplished by using the CEEB Standard Rating and the COOP Converted Score, transforming each into a new standard score on a scale having a mean of 50 and a standard deviation of 10, and adding the two resulting transformed scores. The computer was instructed to examine the scores of an experimental student and to search the control pool for the best possible match. As indicated in the discussion above, the ratio between the experimental pool and the control pool was approximately one to two.

Evaluative Instruments

Three tests of performance in composition were used: the Cooperative English Tests: English Expression (COOP), the College Entrance Examination Board English Composition Test (CEEB), and a theme.

Objective tests. COOP and CEEB are objective tests. The COOP appealed to the investigators because it had been employed in previous research at the University of Northern Iowa and seemed to serve as a reasonably satisfactory indirect measure of student writing ability. The CEEB, unlike the COOP, is a "secure" test. It is changed from administration to administration and a serious attempt is made to assure that students will have no prior access to any of the test items. It was included in part because of its greater security and in part because of a high correlation which had on one occasion been secured between performance on it and evaluations of writing samples. Following is a list of the specific test forms employed on the successive testing occasions:

Testing Date	COOP	CEEB
September 1964	Forms A & B	GB03
January 1965	alternating	HB01
May 1965	at each	HB02
May 1966	university	JB02

The COOP contains 90 items--30 on Effectiveness and 60 on Mechanics. Total time limit is 40 minutes. The CEEB contains from 100 to 110 items and has a total working time of 60 minutes--20 minutes recommended for each of three sections. From test form to test form the elements tested by the CEEB vary somewhat. Representative elements include paragraph organization, construction shifts, sentence correctness, and usage. The various forms of the test are regarded as equivalent but not parallel.

Theme. The theme was a paper written within a two-hour period on a single topic provided by the investigators. Students were urged to remain for the full two-hour period, though they were allowed to leave after an hour and twenty minutes. An explanation of the method for selecting topics, a theme instruction sheet, and the topics used on the various testing dates are included as Appendix A.

Themes were evaluated by teams selected by Fred Godshalk, Chairman of Test Development in the Humanities at the Educational Testing Service, from the pool of readers used by the Educational Testing Service in its theme-reading program. These teams were used because of their wide experience with theme reading and because many of the same readers would be used on successive scoring occasions.

The ETS readers were accustomed to a 4-point scale. The University of Northern Iowa investigators preferred a 9-point scale. The goal was to employ a scoring scale which would permit the separation of the themes into a reasonable number of quality levels without presenting the evaluators with so many rating categories that undue time would be consumed in pondering fine distinctions. A compromise was adopted: a 9-point scale (1 to 9) with emphasis on 2, 4, 6, and 8.

When Mr. Godshalk communicated his standards to the readers, they were asked to think of the normal curve as split in the middle, with each segment so created split again halfway between the median and the extreme. This created four categories: much below average, below average; above average, much above average. It did not provide specifically for the average rank. Readers, already accustomed to the 4-point scale, found it easy to use 2, 4, 6, and 8 as their main grades, but they were able also to use the odd numbers whenever it seemed that a particular paper had some characteristic requiring a grade between two of the even numbers. Since each paper was read by two readers and the ratings summed, the total possible range of scores for a single paper was from 2 to 18. An explanation of the reading procedure is given in Appendix C.

It is recognized that the validity of these evaluations depends upon the degree to which Mr. Godshalk's judgment of student writing, as modified by discussion with the readers, is sound. Mr. Godshalk has an unusually wide background in evaluating the writing of college-bound high school seniors.[2] The readers were from a variety of geographical backgrounds and a wide range of educational institutions. Mr. Godshalk has for years supervised groups of readers like these; the readers have worked together as teams in just such reading situations. Though neither Mr. Cowley nor Mr. Jewell consistently compared their evaluation of sample themes with that of the groups, when they did, there was no pronounced disparity between their ratings and those of the readers. In the judgment of the investigators, the validity of theme evaluations is as high as it is possible to achieve in a project of this sort.

---

[2] Godshalk, Fred, Frances Swineford, and William E. Coffman. The Measurement of Writing Ability, New York: College Entrance Examination Board, 1966.

A different topic was used on each testing occasion, the themes were evaluated at different times, and reader personnel shifted from reading to reading. For these reasons, to report differences between testing dates as gains would be misleading, and this has not been done. Rather, differences between the experimental subgroup and control subgroup after the exact matching in September 1964 are presumed to result from the absence or presence of instruction.

Reliability

Cooperative English Tests: English Expression. This instrument, published in 1960, is composed of two parts: "Part I: Effectiveness," thirty items; and "Part II: Mechanics," sixty items. The time limits are 15 minutes and 25 minutes respectively. A student's score is the total number of correct responses. This raw score is transformed into a Converted Score by means of a table provided by the publishers of the test. For Form 1A, the possible range in converted scores is from 115 (raw score of 0) to 191 (raw score of 90). For the two forms of the test (1A, 1B) recommended for use with college freshmen and sophomores, the investigators were able to find reliability evidence only for the twelfth grade level. The correlation between parallel forms was 0.84 and the standard error of measurement was on the order of 4.00 converted score units.

The College Entrance Examination Board English Composition Test. This is one of the CEEB achievement tests. Evidence about the functioning of this instrument seems to be directly concerned with validity. This is reflected in one of the earlier reports on the instrument, which appeared with the title "Composition Test Shows High Validity on Reliable Criterion of Writing Ability."[3] The excellent 84-page report called The Measurement of Writing Ability[4] also dealt primarily with the validity of the College Entrance Examination Board English Composition Test (CEEB). It is realized that to achieve validity a test author must at the same time achieve reliability. A third source of information was The Sixth Mental Measurements Yearbook. Holland Roberts, one of the three reviewers of the test, commented on reliability: "For the composition test a Kuder-Richardson formula 20 reliability of 0.85 and a standard error of measurement of 39 is reported, indicating satisfactory discrimination among the members of the test group."[5]

---

[3] "Composition Test Shows High Validity on Reliable Criterion of Writing Ability," ETS Developments, XI (January 1963) 1 & 4.

[4] Godshalk, op. cit.

[5] Roberts, Holland [a review of the CEEB English Composition Test], Sixth Mental Measurements Yearbook. Ed. Oscar K. Buros. Highland Park, New Jersey: Gryphon Press, 1965, p. 590.

Theme. The theme test consisted of an impromptu paper, 300-500 words in length, written within a two-hour period. A new topic was used at each testing session, and at each session only one topic was provided. Typically, the topic consisted of a quotation set in a framework intended to link the topic and the student's experience (see Appendix A). Experimental and control students wrote at the same time and in the same place.

Each theme was evaluated by two readers working independently (see discussion, page 41). Each reader assigned each paper a numerical value on a scale extending from 1 to 9. It is thus possible to examine the extent of between-reader agreement in assigned ratings.

As stated in the Interim Report, page 55, the investigators believe that a meaningful basis for thinking about theme reliability is in terms of the extent of agreement between the two independent ratings of each theme. For the present discussion, the theme scores of 90 matched pairs of students were examined. The 90 matched pairs were all of the pairs available at one university in May 1965; they are included among the 365 matched pairs, the total for four institutions,* whose theme performance of May 1965 is reported in Table XXVI. The tabulation below displays the inter-reader consistency in theme ratings for the 180 students--90 experimentals plus 90 controls.

Difference in Two Ratings of Theme	Number of Themes
0	46
1	62
2	52
3	9
4	9
5	2

Mean Difference = 1.33

This tabulation gives only the absolute value of the differences. An estimate of a reliability coefficient derived from a distribution of differences in the two assigned ratings would need to take into account the direction as well as the amount of the differences.

---

*After the first semester the data at institution 4 were limited and incomplete and therefore not included in the totals.

Approximately one-fourth of the 180 themes were assigned the same rating by the two independent readers. More than half of the 180 themes (108) were rated no more than 1 point apart. Only 20 of the 180 papers showed an inter-reader discrepancy of more than 2 points. The maximum inter-reader discrepancy was 5 (for 2 of 180 themes). The maximum possible inter-reader discrepancy was 8.

The degree of inter-reader consistency portrayed in the above tabulation can also be represented by a coefficient of correlation. The tabulation below indicates that the Pearson product-moment r between the two sets of ratings for the 90 experimental themes was 0.22, for the 90 control themes 0.18. Thus if the theme were regarded as a 9-point test, the coefficient of correlation of rating consistency would be on the order of 0.20.

			Reader 1		Reader 2		Reader 1 + 2	
Subgroup	N	r	Mean	S.D.	Mean	S.D.	Mean	S.D.
Experimental	90	0.22	4.98	1.30	4.83	1.34	9.81	2.05
Control	90	0.18	4.92	1.34	5.03	1.31	9.96	2.04

However, it is more appropriate to regard the theme as an 18-point test, for the score used for each student was the sum of the two ratings, with a potential range of 2 to 18. The coefficient of 0.20 could then be conceived as the correlation between scores on two readings of a half-test. Actually, however, it is the range of the rating scale rather than the length of the test which is being doubled. In such a context it is possible to estimate the inter-reader correlation on an 18-point scale by basing the correlation on two ratings of the same test. The Spearman-Brown Prophecy Formula would yield a coefficient of 0.33 in this event.

The foregoing discussion of theme rating reliability in terms of coefficients of correlation suggests the complexity of the assumptions, interpretations, and arbitrary decisions involved. For most purposes the extent of agreement between two independent ratings assigned to a single paper, as illustrated above for 180 papers, provides the clearest picture of theme-rating reliability.

Inter-reader agreement represents only one aspect of theme reliability. Involved also is the fact that a student's performance probably differs from day to day and from topic to topic. Because of problems like these, though there is considerable acceptance of a theme as a desirable form of measuring instrument in English composition, there is considerable doubt that ratings can be assigned reliably. Thus, degree of inter-reader consistency in assigned ratings is of major interest. While the grading of any essay test is difficult, the grading of themes is especially complex, particularly because there is a minimum of commonality among satisfactory responses to a topic.

FINDINGS AND ANALYSIS

General Data

The data are treated first to show the nature of the persisting samples and then to clarify the results at each testing period on each criterion measure. Testing was done on four occasions--beginning of the first semester, end of the first semester, end of the second semester, end of the fourth semester. The numbers of matched pairs were, respectively, 1,040 (original group), 597, 365, and 122.

The Samples

Table I presents a composite picture of the entering freshmen at the participating universities, reporting their performance in September 1964, on seven variables. The number of freshmen per institution varied from 705 to 943, selected in the manner described on page 38. The 4,190 freshmen, representative of the freshmen entering the five participating universities in September 1964, constitute the sample from which all subgroups were drawn. The data in Table I provide evidence of the extent to which the persisting experimental and control subgroups, composed of matched pairs of students, remain representative of the parent group. None of the information in Table I involves student performance after September 1964.

Line one shows the performance of the 4,190 students--the experimental pool plus the control pool--in September 1964. For example, their mean percentile rank in high school class was 67.02.*

---

*The percentile rank data were reported by four of the participating institutions as the percentage of individuals with a high school rank lower than that of the given individual. One institution reported the tenth in which each individual ranked; in this instance the investigators used 95, 85, 75, ... We have computed mean percentile rank, realizing the limitations of such a procedure.

TABLE I

ACHIEVEMENT AS OF SEPTEMBER 1964 OF THE SAMPLE OF 1964-65 FRESHMAN STUDENTS AND OF VARIOUS PERSISTING PORTIONS OF THAT SAMPLE: COMBINED INSTITUTIONS

No. of Institutions	Sample or Subgroup	N	Per-cent Men	Percentile Rank in H. S. Class Mean	Percentile Rank in H. S. Class S. D.	N	COOP Eng. Exp. Converted Score Mean	COOP Eng. Exp. Converted Score S. D.	N	CEEB Eng. Comp. Standard Rating Mean	CEEB Eng. Comp. Standard Rating S. D.
5	Exp. Pool plus Control Pool	4,190	42.7	67.02	19.63	4,152	161.68	8.73	4,159	472.85	96.14
5	Exp. Pool	1,408	43.9	66.37	19.73	1,399	161.97	8.64	1,399	473.52	96.16
5	Control Pool	2,782	42.1	67.35	19.57	2,753	161.58	8.77	2,760	472.51	96.13
5	Matched Exp. Sept. 1964	1,040	40.6	66.33	19.37	1,040	161.59	7.69	1,040	472.04	86.16
5	Matched Cont. Sept. 1964	1,040	40.6	67.92	18.93	1,040	161.63	7.61	1,040	472.13	85.58
5	Experimentals January 1965	597	39.4	69.31	18.03	597	162.29	7.39	597	477.75	84.64
5	Controls January 1965	597	39.4	70.44	18.39	597	162.24	7.50	597	478.40	83.41
4	Experimentals May 1965	365	36.7	70.82	17.40	365	162.60	7.30	365	484.19	83.25
4	Controls May 1965	365	36.7	71.48	18.40	365	162.79	7.40	365	482.81	82.85
4	Experimentals May 1966	122	36.1	73.93	18.40	122	163.44	7.45	122	496.88	85.67
4	Controls May 1966	122	36.1	71.64	20.06	122	163.85	7.23	122	493.32	86.07

TABLE I
CONTINUED

No. of Institu- tions	Sample or Subgroup	N	Z-Score[1] Mean	Z-Score[1] S. D.	N	Theme Rating 1 Mean	Theme Rating 1 S. D.	September Theme Theme Rating 2 Mean	September Theme Theme Rating 2 S. D.	Theme Rating Total Mean	Theme Rating Total S. D.
5	Exp. Pool plus Control Pool	4,136	[illegible].49	18.45	4,148	4.62	1.57	4.41	1.52	9.03	2.59
5	Exp. Pool	1,392	99.70	18.41	1,396	4.64	1.59	4.43	1.56	9.07	2.49
5	Control Pool	2,744	99.38	18.47	2,752	4.61	1.56	4.40	1.50	9.01	2.57
5	Matched Exp. Sept. 1964	1,040	99.31	16.04	1,040	4.62	1.44	4.38	1.40	9.00	2.22
5	Matched Cont. Sept. 1964	1,040	99.36	15.92	1,040	4.61	1.41	4.38	1.36	9.00	2.22
5	Experimentals January 1965	597	99.93	15.69	597	4.70	1.38	4.45	1.41	9.15	2.14
5	Controls January 1965	597	99.95	15.59	597	4.71	1.38	4.43	1.33	9.15	2.14
4	Experimentals May 1965	365	100.88	15.34	365	4.76	1.30	4.48	1.41	9.23	2.01
4	Controls May 1965	365	100.96	15.16	365	4.76	1.35	4.47	1.25	9.23	2.01
4	Experimentals May 1966	122	103.49	15.58	122	4.68	1.20	4.63	1.37	9.31	1.87
4	Controls May 1966	122	103.60	15.33	122	4.80	1.27	4.51	1.16	9.31	1.87

---

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test September scores. See page 39

Line two indicates that the 1,408 members of the experimental pool had a mean percentile rank in high school class of 66.37, while line three indicates that the 2,782 members of the control pool had a mean percentile rank in high school class of 67.35. On the remaining six variables, means for the experimental pool and the control pool also show a close similarity.

Establishment of matched pairs resulted in 1,040 members of the experimental pool being paired with members of the control pool. The 1,040 matched experimentals made up the experimental subgroup; the 1,040 matched controls made up the control subgroup. The fact that matching was exact for Total Theme Rating may be seen in the means of 9.00 and standard deviations of 2.22. The subgroup means were also similar on the other variables. For example, mean percentile rank in high school graduating class was 66.33 for the experimentals and 67.92 for the controls. It is worthy of note that the means for the matched experimentals and the matched controls were close to the means for the respective pools. That is, the process of forming matched pairs yielded experimental and control subgroups representative of the parent group--the 4,190 entering freshmen who constituted the project pool.

Attrition reduced the number of matched pairs from 1,040 in September to 597 in January. In the matched pairs design a complete matched pair must be dropped if only one member of the pair leaves.* The degree to which the two subgroups have been "refined" by the loss of members over the first semester may be examined. Comparing line four with line six reveals that the 1,040 members of the experimental subgroup beginning the semester had a mean percentile rank in high school class of 66.33, while the 597 members of the subgroup completing the first semester had a mean percentile rank in high school class of 69.31. The corresponding facts for the control subgroup are 67.92 and 70.44. Analyses for the other variables show that there was a similar selectivity factor operating which caused the January subgroups to be slightly superior to the larger parent subgroups.

Data for the full freshman year also show the selectivity associated with attrition. At the end of the academic year, the percentile rank for the experimental subgroup (N=365) was 70.82, while that of the control subgroup was 71.48--changes of 4.49 for the experimental subgroup and 3.56 for the control subgroup.

---

*Though the matched pairs design is vulnerable to high attrition, it has advantages which counterbalance this weakness. See discussion, Appendix B.

Data for the complete two-year period likewise show the influence of attrition and absences from test sessions. In May 1966 the persisting members of the experimental subgroup (N=122) show a mean percentile rank in high school class of 73.93, and the control 71.64.

Another way of examining the extent to which persisting subgroups of matched pairs excelled the original subgroups of matched pairs is in terms of the placement of mean scores in the September distribution of student scores. For this purpose the CEEB test and the distribution of scores for the 4,159 students who, at the outset, comprised the experimental pool plus the control pool will be used. The tabulation below shows that, for the 1,040 matched experimental students, the mean CEEB English Composition Standard Rating was 472.04 and that for the control students the mean was 472.13. Each of these means lies at approximately the 49th percentile rank in the distribution of the 4,159 scores.

Mean CEEB Standard Rating September 1964	Subgroup	Percentile Rank Based on September 1964 Distribution (N=4,159)
472.04	September 1964 Experimentals (N=1,040)	49
472.13	September 1964 Controls (N=1,040)	49
484.19	May 1965 Experimentals (N=365)	54
482.81	May 1965 Controls (N=365)	53
496.88	May 1966 Experimentals (N=122)	58
493.32	May 1966 Controls (N=122)	57

This analysis has shown that whereas among the 1,040 original matched pairs the typical CEEB score had a percentile rank of 49 in September (close to the expected 50), for the 122 matched pairs who completed testing through two years of college the typical CEEB score had a percentile rank of about 58 in September. Thus the degree of selectivity over the two years was on the order of 8 to 10 percentile rank points on the CEEB.

At the end of the freshman year (May 1965), 365 matched pairs remained. The experimental subgroup had a September CEEB mean of 484.19, which corresponds to a percentile rank of 54 in the distribution of the 4,159 scores. The corresponding figures for the control subgroup were 482.81, and 53.

The tabulation above also shows the percentile rank in the September 1964 student score distribution for the mean scores of the members of experimental and control subgroups who completed the May 1966 testing. For the 122 experimentals the September CEEB mean was 496.88, and the percentile rank 58. The 122 controls had a mean of 493.32, and a percentile rank of 57.

All Tests--September 1964 through May 1966

Whereas in Table I all test scores were those available in September 1964, both Table II and Table III present performance at successive testing occasions. Table II presents the performance at each of the four successive testing periods beginning with September 1964, of the 122 matched pairs who completed the entire testing program; Table III portrays the performance of all persisting matched pairs at each of these four successive testing periods. Inspection of this table will reveal the differences in performance on each of the criterion measures for the two subgroups at the beginning of the fall semester, 1964-65; at the end of the fall semester, 1964-65; at the end of the spring semester, 1964-65; and at the end of the spring semester, 1965-66.

The experimental students did not receive instruction in freshman composition; the control students did. The data in Table III permit the key comparisons of the project; those between the performance of the experimental and control subgroups on the criterion measures at successive points in their college careers.

Intercorrelation Data

Table IV shows product-moment coefficients of correlation for all possible pairs of variables among a set of eight variables. The 597 students are the experimental members of matched pairs who completed the first semester of the freshman year at the participating institutions.

The following specific points may be noted:

How did percentile rank in high school class correlate with all the measures of English composition ability? Between 0.20 and 0.30.

TABLE II

PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON EACH OF THREE CRITERION MEASURES AT BEGINNING, MIDDLE, AND END OF FIRST YEAR AND END OF SECOND YEAR OF COLLEGE: COMBINED INSTITUTIONS

		Cooperative English Tests: English Expression (1960) Converted Score		CEEB English Composition Test Standard Rating		Theme Rating 1		Theme Rating 2		Theme Rating Total	
Subgroup	Time of Testing	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.
Experimental	Sept. 1964	163.44	7.45	496.88	85.67	4.68	1.20	4.63	1.37	9.31	1.87
Control	"	163.85	7.23	493.32	86.07	4.80	1.27	4.51	1.16	9.31	1.87
Experimental	Jan. 1965	165.32	7.06	531.74	81.49	4.97	1.42	5.24	1.40	10.20	2.24
Control	"	166.44	7.09	516.54	90.76	5.24	1.30	5.18	1.41	10.42	2.08
Experimental	May 1965	167.91	6.79	531.77	74.62	5.06	1.36	4.88	1.38	9.93	2.25
Control	"	168.76	6.79	536.08	78.77	4.92	1.38	4.95	1.33	9.87	2.14
Experimental	May 1966	167.24	8.93	554.62	81.65	4.82	1.40	4.89	1.57	9.70	2.63
Control	"	168.52	7.89	552.11	75.96	4.74	1.49	4.65	1.46	9.39	2.59

NOTE: The small, but perturbing, problems with the rounding of decimal values were encountered with portions of this table. The investigators recognize that in practically every table in this report there are spots at which the "rounding dilemma" may have produced slight inconsistencies in the reported values.

TABLE III

PERFORMANCE OF AVAILABLE SAMPLES OF MATCHED PAIRS OF STUDENTS ON THREE CRITERION MEASURES AT BEGINNING, MIDDLE, AND END OF FIRST YEAR AND END OF SECOND YEAR OF COLLEGE: COMBINED INSTITUTIONS

			Cooperative English Tests: English Expression (1960) Converted Score		CEEB English Composition Test Standard Rating		Theme Rating 1		Theme Rating 2		Theme Rating Total	
Subgroup	N	Time of Testing	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.
Experimental	1,040	Sept.1964	161.59	7.69	472.04	86.16	4.62	1.44	4.38	1.40	9.00	2.22
Control	1,040	"	161.63	7.61	472.13	85.58	4.61	1.41	4.38	1.36	9.00	2.22
Experimental	597	Jan. 1965	163.93	7.83	507.12	81.83	4.81	1.49	4.88	1.49	9.69	2.47
Control	597	"	164.94	7.63	508.88	82.24	4.97	1.46	4.99	1.54	9.96	2.45
Experimental	365	May 1965	166.50	6.82	516.98	82.01	4.83	1.36	4.80	1.36	9.63	2.27
Control	365	"	167.29	7.09	524.28	78.94	4.82	1.39	4.93	1.39	9.76	2.21
Experimental	122	May 1966	167.24	8.93	554.62	81.65	4.82	1.40	4.89	1.57	9.70	2.63
Control	122	"	168.52	7.89	552.11	75.96	4.74	1.49	4.65	1.46	9.39	2.59

TABLE IV

INTERCORRELATIONS AMONG EIGHT VARIABLES FOR 597 EXPERIMENTAL STUDENTS IN SEPTEMBER 1964 AND JANUARY 1965: COMBINED INSTITUTIONS

Variable	Percentile Rank in H. S. Class	Z-Score[1]	COOP Eng. Conv. Score Sept. 1964	COOP Eng. Conv. Score Jan. 1965	CEEB Eng. Comp. Sept. 1964	CEEB Eng. Comp. Jan. 1965	Theme Rating Sept. 1964	Theme Rating Jan. 1965
Percentile Rank in H. S. Class								
Z-Score (COOP English plus CEEB)	0.28							
COOP English Test Converted Score September 1964	0.24	0.89						
COOP English Test Converted Score January 1965	0.30	0.67	0.66					
CEEB Stan. Rating September 1964	0.27	0.88	0.64	0.60				
CEEB Stan. Rating January 1965	0.27	0.71	0.63	0.56	0.65			
Theme Rating September 1964	0.23	0.47	0.43	0.39	0.47	0.39		
Theme Rating January 1965	0.20	0.46	0.46	0.40	0.44	0.42	0.47	

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

How did the September scores on the two objective tests correlate with the January theme scores? 0.46 for COOP; 0.44 for CEEB; 0.46 for COOP + CEEB (Z-score).

How did the January scores for the two objective tests correlate? 0.56 for COOP vs. CEEB.

How did the September scores and January scores correlate? 0.66 for COOP; 0.65 for CEEB; 0.47 for Theme. How did the September COOP and CEEB scores correlate? 0.64.

Did the January scores on the CEEB correlate higher with the September CEEB scores than with the September COOP scores? Possibly slightly; 0.65 with CEEB, 0.63 with COOP.

The above intercorrelation data for the experimental students are similar enough to the data for the control members of the matched pairs which are reported in Table V that no detailed discussion of Table V is included. In general, the magnitude of the correlation coefficients is in line with those for other similar situations, including the *Interim Report* of the present study.

First Semester Sample

COOP, September 1964-January 1965

Table VI is the first in the series of tables in which, for each of the three measuring instruments, student performance is analyzed to show basic comparisons of test performance for persisting experimental and control students: within subgroups between beginning and final means, and between subgroups. Associated means, standard deviations, r's, t's, and male and female comparisons are also displayed in these tables.

Overall performance. The primary comparison in Table VI is between the means for the experimental subgroup and the control subgroup after the first semester of college (1964-65). On COOP the difference in means was 1.01--164.94 for the controls and 163.93 for the experimentals. The correlation between the scores of the 597 matched pairs of students was 0.52. The t-value of 3.273 is significant. It is noteworthy that a relatively small difference in means--just over one converted score point--is significant.

TABLE V

INTERCORRELATIONS AMONG EIGHT VARIABLES FOR 597 CONTROL STUDENTS IN SEPTEMBER 1964 AND JANUARY 1965: COMBINED INSTITUTIONS

Variable	Percentile Rank in H. S. Class	Z-Score[1]	COOP Eng. Conv. Score Sept. 1964	COOP Eng. Conv. Score Jan. 1965	CEEB Eng. Comp. Sept. 1964	CEEB Eng. Comp. Jan. 1965	Theme Rating Sept. 1964	Theme Rating Jan. 1965
Percentile Rank in H. S. Class								
Z-Score (COOP English plus CEEB)	0.32							
COOP English Test Converted Score September 1964	0.27	0.88						
COOP English Test Converted Score January 1965	0.32	0.72	0.69					
CEEB Stan. Rating September 1964	0.32	0.88	0.62	0.65				
CEEB Stan. Rating January 1965	0.23	0.63	0.54	0.62	0.62			
Theme Rating September 1964	0.26	0.47	0.44	0.39	0.46	0.39		
Theme Rating January 1965	0.19	0.40	0.36	0.43	0.37	0.33	0.37	

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

TABLE VI

THE PERFORMANCE OF 597 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

Subgroup	N	Cooperative Eng. Test Converted Score				Difference in Means, Jan. minus September	r	t-Ratio	Difference in Sept. Means Control minus Experimental			Difference in Jan. Means Control minus Experimental			Degrees of Freedom
		September		January											
		Mean	S. D.	Mean	S. D.				Diff.	r	t	Diff.	r	t	
Experimental	597	162.29	7.39	163.93	7.83	1.64	0.66	6.367*	0.05	0.81	0.266	1.01	0.52	3.273*	596
Control	597	162.24	7.50	164.94	7.63	2.70	0.69	10.999*							
Exp. Males	235	160.00	7.35	161.64	8.20	1.64	0.55	3.382*	-0.14	0.78	0.431	1.06	0,49	2.053*	234
Cont.Males	235	159.86	7.72	162.70	7.43	2.84	0.64	6.792*							
Exp. Females	362	163.77	7.03	165.41	7.20	1.64	0.71	5.733*	0.01	0.81	0.036	0.98	0.50	2.549*	361
Cont.Females	362	163.78	6.93	166.39	7.40	2.61	0.68	8.675*							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	3.77	6.276*	3.77	5.912*	595
Controls	3.92	6.436*	3.69	5.935*	595

*Significant at 0.05 level (two-tailed test);

One of the factors in this result is the relatively large sample (N=597) which yielded a relatively small standard error (S.E.=0.31). Relationships between size of sample and size of standard error may be seen in data from the following tabulation which is calculated from data in the Interim Report which preceded the present report.[6]

Subgroups	Number of Matched Pairs	Difference in Means	Standard Error of the Difference
End-of-first semester	166	0.59	0.51
End-of-second semester	113	0.66	0.75
End-of-fourth semester	31	0.23	1.28

In Table VI the experimental-control mean difference in January is shown as 1.01; the t-value is 3.273, which is significant. The standard error is 0.309. This indicates superiority for the control students--those who had received a semester of composition instruction.

These figures, when compared to those reported in a relevant section of the Interim Report, indicate the value of relatively large samples when differences between means are small.[7] In the pilot study, with only 31 matched pairs, the difference in means on the COOP between the experimental subgroup and the control subgroup in January of the first year was 3.00, the standard error 1.13, and the t-ratio 2.65. Had the standard error for the 597 matched pairs in the current study been 1.13, the t-value would have been 0.893, which would not have been significant, instead of 3.273, which is significant.

This superiority for the control students over the experimental students was found to be about the same for males as for females. It will be seen from Table VI that the differences in January means were 1.06 for males, and 0.98 for females. Both are significant.

---

[6] Interim Report, Table IX, page 36.

[7] Ibid.

The lower portion of Table VI compares males and females of the experimental subgroup as well as males and females of the control subgroup. In January the females had mean scores about 3.70 higher than the males in both the experimental and control subgroups. This superiority of the females is significant.

It is instructive to examine the mean change scores from September to January. The mean gains varied from 1.64 to 2.84 for the two principal subgroups and their male and female components. What is the meaning in terms of test performance of a change on the order of two COOP Converted Score points over a semester? If one works within the September distribution of test scores and their associated percentile ranks, it is evident that an increase of two Converted Score points may be achieved by an increase of three raw score points. Advances of this magnitude result in corresponding percentile rank increase of 10 or 11 for scores near the median and of 4 for scores having percentile ranks of 10 and 90.

The most noteworthy finding presented in Table VI is the extent to which females as a subgroup excel males as a subgroup. The superiority of 3.77 or 3.69 is almost twice as large as the change in mean scores for the entire group during one semester of college. Thus at the beginning of the first semester, females possess a higher mean test score than the males do at the end of the first semester.

Performance by ability quarters. In Table VII the emphasis is on performance of experimental students and control students at each of four ability levels. The ability levels were established on the basis of student performance in September 1964 on two tests: COOP and CEEB. A Z-score was obtained by combining derived standard scores for these tests.* The four ability levels were based on the 4,136 students for whom Z-scores were available: the experimental pool plus the control pool for all five institutions. The fact that the highest quarter and the lowest quarter contain the smallest numbers of students (N=126 and N=134) may be explained primarily by the difficulties in matching near the extremes of the distribution.

Evidence presented in Table VII enables one to answer the following question: Was the superiority of the control students about equally present at all ability levels? There is considerable fluctuation, the second highest quarter showing the greatest control-experimental difference (1.91) and the third highest quarter least

---

*See discussion of Z-score, page 39.

TABLE VII

THE PERFORMANCE OF 597 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965, BY Z-SCORE[1] QUARTERS: MEANS, STANDARD DEVIATIONS, AND t-RATIOS: COMBINED INSTITUTIONS

			Sept. 1964		Jan. 1965		Difference in Means, January 1965 minus September 1964			Difference in January Means, Control minus Exper.			
Quarter	Subgroup	N	Mean	S. D.	Mean	S. D.	Difference	r	t-Ratio	Difference	r	t-Ratio	Degrees of Freedom
Highest 1/4	Experimental	126	171.28	4.00	171.63	6.36	0.36	0.41	0.669	0.95	0.36	1.581	125
"	Control	126	170.97	4.23	172.59	5.45	1.62	0.52	3.712*				
Second High 1/4	Experimental	178	164.56	3.42	165.24	6.70	0.69	0.26	1.367	1.91	0.08	2.974*	177
"	Control	178	164.84	3.81	167.15	5.91	2.31	0.25	4.965*				
Third High 1/4	Experimental	159	160.03	4.15	162.04	5.30	2.01	0.34	4.572*	0.31	0.01	0.547	158
"	Control	159	160.12	3.81	162.36	4.97	2.24	0.27	5.224*				
Lowest 1/4	Experimental	134	153.51	5.13	157.16	5.81	3.66	0.34	6.707*	0.71	0.27	1.143	133
"	Control	134	153.09	5.27	157.87	6.01	4.78	0.27	8.092*				
Total Group	Experimental	597	162.29	7.39	163.93	7.83	1.64	0.66	6.367*	1.01	0.52	3.273*	596
"	Control	597	162.24	7.50	164.94	7.63	2.70	0.69	10.999*				

[1]

*Significant at 0.05 level (two-tailed test).

(0.31). However, at all four ability levels the January mean for the members of the control subgroup was higher than the mean of the experimental subgroup.

This kind of review of the evidence by ability levels as well as for the overall subgroups will be reported for each of the first two semesters and for the full freshman year.

Since for the total sample of 597 matched pairs the control mean was significantly higher than the experimental mean, it would be anticipated that there would be a statistically significant superiority in favor of the control subgroup at some of the ability levels. It was only in the second highest quarter that this was found. Such a straightforward analysis, although an over-simplification of what would be required in a thorough consideration of relationships between main effects and interaction, is nevertheless useful in examining the extent to which findings for the various ability levels seemed to be consistent with the findings for the total subgroups.

Performance by ability quarters by sex. Before proceeding to a consideration of the evidence for the two sexes at each of four ability levels, it is important to recognize a feature of the situation which has a distinct bearing on the analysis.

A positive correlation exists between sex and performance on tests of English ability with females outperforming males (Tables VI and VIII). This results in the presence (both in the initial and the surviving matched pairs) of a larger female-male ratio in the top ability levels than in the bottom ability levels. In the top quarter (Table VIII) the ratio is 99 to 27 (3.6 to 1) and in the lowest quarter the ratio is 56 to 78 (0.72 to 1). The mean differences by quarters between females and males tend to be systematically smaller than the overall mean difference between the sexes. The reader may note this fact by comparing the mean difference in January COOP of 3.77 between experimental females and experimental males overall (Table VI) with the mean female-male differences at the four ability levels as presented in Table VIII: 1.56, 1.38, 0.99, and 1.66. To arrive at the overall mean difference by using the mean differences at each of the four ability levels, it would be necessary to employ a weighting procedure which took into account the fluctuation of the female-male ratios at the four ability levels. This characteristic of the data, while a proper reflection of the samples used, does introduce a complication in interpretation of the evidence presented in the quarters-by-sex tables (Tables VIII, XI, XIV, and XXIX).

Table VIII completes a series which reports COOP evidence for the first semester of the freshman year (1964-65). The uniqueness of Table VIII is in the presentation of the facts for males and

√

TABLE VIII

PERFORMANCE ON COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION BY MATCHED PAIRS BY SEX IN SEPTEMBER 1964 AND JANUARY 1965 AT EACH OF FOUR Z-SCORE[1] LEVELS: COMBINED INSTITUTIONS

Subgroup	Level	N	Sept. Z-Score Rating Mean	S.D.	Sept. COOP Conv. Score Mean	S.D.	Jan. COOP Conv. Score Mean	S.D.	COOP Jan. minus Sept. Mean	r	t-Ratio	df	COOP, Female minus Male September Diff.	t-Ratio	January Diff.	t-Ratio	df
Exp.Male	Highest	27	117.93	4.41	169.81	4.13	170.41	8.28	0.59	0.17	0.351	26	1.87	2.169*	1.56	1.128	124
Exp.Female	1/4	99	122.73	7.06	171.68	3.86	171.97	5.68	0.29	0.50	0.579	98					
Cont.Male	"	27	117.74	3.90	169.30	3.21	170.74	4.69	1.44	0.63	2.026*	26	2.12	2.348*	2.35	2.001*	124
Cont.Female		99	122.53	6.93	171.42	4.36	173.09	5.54	1.67	0.48	3.196*	98					
Exp.Male	2nd	65	105.75	3.58	164.74	3.43	164.36	8.22	-0.37	0.20	0.357	64	-0.29	0.537	1.38	1.317	176
Exp.Female	High 1/4	113	105.50	3.43	164.45	3.40	165.74	5.57	1.29	0.34	2.505*	112					
Cont.Male	"	65	105.66	3.86	165.00	3.73	166.94	5.68	1.94	0.14	2.447*	64	-0.25	0.416	0.33	0.363	176
Cont.Female		113	105.61	3.60	164.75	3.85	167.27	6.03	2.52	0.30	4.383*	112					
Exp.Male	3rd	65	93.88	3.74	159.63	4.43	161.46	5.60	1.83	0.26	2.399*	64	0.68	1.008	0.99	1.149	157
Exp.Female	High 1/4	94	93.80	3.50	160.31	4.02	162.45	5.05	2.14	0.39	4.046*	93					
Cont.Male	"	65	93.65	3.86	160.37	4.15	162.15	4.64	1.78	0.25	2.644*	64	-0.42	0.684	0.35	0.429	157
Cont.Female		94	93.93	3.38	159.95	3.55	162.50	5.19	2.55	0.29	4.597*	93					
Exp.Male	Lowest	78	77.24	8.60	152.97	5.47	156.47	6.05	3.50	0.35	4.645*	77	1.28	1.419	1.66	1.626	132
Exp.Female	1/4	56	81.86	4.72	154.25	4.51	158.13	5.30	3.87	0.31	4.954*	55					
Cont.Male	"	78	77.44	8.89	151.90	5.81	156.85	6.49	4.95	0.29	5.907*	77	2.85	3.180*	2.45	2.366*	132
Cont.Female		56	82.18	4.91	154.75	3.85	159.30	4.92	4.55	0.08	5.640*	55					

---

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

*Significant at 0.05 level (two-tailed test).

females by ability level. In the highest quarter, within the experimental subgroup, the January mean for the 99 females was 1.56 higher than the mean for the 27 males. This difference is of about the order of the difference which prevailed on the September scores (1.87). Among the control students in the highest quarter, the females had a January mean which was 2.35 higher than the male mean. In September the difference had been 2.12.

In addition to the top quarter, the lowest quarter also revealed noticeable superiority of females over males on the COOP. Within the experimental subgroup the mean difference in January was 1.66; within the control subgroup, 2.45.

In the two middle quarters, there was a consistent but smaller advantage in favor of the females on January scores. It is noteworthy, however, that on the September scores, three of the four comparisons show the males to be slightly ahead of the females.

The data in Table VIII illustrate a fact which the investigators have emphasized: females out-perform males as groups, appreciably and consistently, even at various ability levels. An inspection of the basic data shows that only at the top edge of the distribution (the top 2 percent) do males equal or excel females. At the bottom edge of the distribution the reverse is true; the male group falls below the females.

CEEB, September 1964-January 1965

Overall performance. Table IX shows that the number of matched pairs completing all September and January CEEB tests was 597. The facts presented are:

(1) Within experimental subgroup and within control subgroup: September mean and standard deviation, January mean and standard deviation, difference in means January minus September; correlation, and t-ratio.

(2) Between experimental subgroup and control subgroup: difference in September means, correlation, t-ratio; difference in January means, correlation, t-ratio.

(3) Data described in (1) and (2), except for correlations, by sex.

For the total experimental subgroup, the January mean was 29.37 higher than the September mean (507.12 minus 477.75). This change in mean is significant; the t-ratio is 10.354 with 596 degrees of freedom. For the total control subgroup, the mean gain was 30.47,

TABLE IX

THE PERFORMANCE OF 597 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in September Means, Control minus Experimental			Diff. in January Means, Control minus Experimental			
		September		January		Jan. minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	597	477.75	84.64	507.12	81.83	29.37	0.65	10.354*	0.66	0.82	0.319	1.76	0.47	0.506	596
Control	597	478.40	83.41	508.88	82.24	30.47	0.62	10.262*							
Exp. Males	235	453.92	85.38	486.45	81.44	32.53	0.65	7.127*	0.68	0.78	0.186	2.47	0.37	0.441	234
Cont.Males	235	454.60	83.40	488.91	70.77	34.32	0.56	7.061*							
Exp. Females	362	493.22	80.46	520.54	79.25	27.32	0.63	7.544*	0.64	0.83	0.263	1.30	0.49	0.293	361
Cont.Females	362	493.86	79.70	521.83	86.47	27.98	0.64	7.471*							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	39.30	5.681*	34.09	5.071*	595
Controls	39.26	5.764*	32.92	4.864*	595

*Significant at 0.05 level (two-tailed test).

also significant (t=10.262, degrees of freedom, 596). Thus the experimental students and the control students advanced about the same amount on CEEB during the fall semester, both gaining significantly.

An analysis of the January test scores reveals that the mean for the control subgroup is 1.76 points higher than the mean for the experimental subgroup. This difference on CEEB is not significant (t=0.506). The data regarding September means for controls and experimentals confirm the similarity achieved in the matching process; the correlation of September CEEB scores between members of matched pairs was 0.82.

The middle portion of Table IX shows the performance of each sex in each of the two subgroups. For the males the mean gains during the semester were 32.53 (experimentals) and 34.32 (controls). The corresponding figures for females were 27.32 (experimentals) and 27.98 (controls). Again, the evidence on CEEB for the first semester shows that the experimental treatment and the control treatment were about equally efficacious in producing change in performance on the CEEB. Mean gains by males exceed mean gains by females by approximately five points. The t-ratios for January-minus-September means are of the order of 7.00 for males and for females.

The lower portion of Table IX shows that within both the experimental subgroup and the control subgroup the mean for females exceeded the mean for males significantly at the beginning and also at the end of the first semester. The mean gains over the semester were slightly higher for the males than for the females--of the order of 33 points compared to 27 points. The superiority of female performance over male performance at a given juncture is of about the same magnitude as the superiority of an end-of-semester mean over a beginning-of-semester mean for either of the sexes. In other words, even among beginning college freshmen, the females are about a semester ahead of the males on the CEEB as a measure of writing ability, and this difference remains at the end of the first semester.

Performance by ability quarters. From Table X it is possible to determine whether the difference of 1.76 in overall means in January CEEB, favoring the 597 control students, resulted from a fairly uniform differential across the four ability levels. Such was not the case. For the lowest one-fourth of the students, the mean for the control subgroup was substantially higher than the mean for the experimental subgroup. For the three highest quarters of ability the means for the control subgroup were slightly lower than the means for the experimental subgroup.

TABLE X

THE PERFORMANCE OF 597 MATCHED PAIRS OF STUDENTS ON THE COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TEST IN SEPTEMBER 1964 AND JANUARY 1965, BY QUARTERS: MEANS, STANDARD DEVIATIONS, AND t-RATIOS: COMBINED INSTITUTIONS

			Sept. 1964		Jan. 1965		Difference in Means, January 1965 minus September 1964			Difference in January Means, Control minus Exper.			
Quarter	Subgroup	N	Mean	S.D.	Mean	S.D.	Difference	r	t-Ratio	Difference	r	t-Ratio	Degrees of Freedom
Highest 1/4	Experimental	126	582.73	46.17	589.97	63.86	7.24	0.34	1.248	- 2.40	0.13	0.299	125
"	Control	126	584.60	45.03	587.56	72.19	2.97	0.22	0.436				
Second High 1/4	Experimental	178	507.26	41.74	526.91	61.27	19.65	0.25	4.033*	- 0.36	0.13	0.054	177
"	Control	178	504.04	42.37	526.55	72.63	22.51	0.32	4.196*				
Third High 1/4	Experimental	159	441.82	46.68	482.92	60.35	41.11	0.21	7.574*	- 5.99	0.04	0.909	158
"	Control	159	440.68	46.58	476.93	59.02	36.25	0.19	6.721*				
Lowest 1/4	Experimental	134	382.46	53.66	431.63	57.67	49.18	0.17	7.888*	17.68	0.15	2.757*	133
"	Control	134	389.26	54.96	449.31	55.68	60.05	0.34	10.922*				
Total Group	Experimental	597	477.75	84.64	507.12	81.83	29.37	0.65	10.354*	1.76	0.47	0.506	596
"	Control	597	478.40	83.41	508.88	82.24	30.47	0.62	10.262*				

*Significant at 0.05 level (two-tailed test).

Another noteworthy feature of Table X is in the column reporting difference in means, January minus September. It will be noted that in general these differences increase from the top ability level to the lowest ability level. For example, in the highest one-quarter the differences are 7.24 (experimental) and 2.97 (control); in the lowest one-quarter the differences are 49.18 (experimental) and 60.05 (control). This inverse relationship between mean gain and ability level, for both the experimental and control subgroups, is consistent with the data reported for COOP in Table VII, page 59, where the top one-quarter showed gains of 0.36 (experimental) and 1.62 (control) while the lowest one-quarter showed gains of 3.66 (experimental) and 4.78 (control). In this kind of an analysis one must recognize the possibility that for a high ability level the increase of scores on a second testing may be restricted by test ceiling. Inspection of our raw data indicates, for example, that the top male on COOP in September had a perfect score, while the top three females were only three points away from perfection. Obviously, on a second testing, their room for improvement was slight. Furthermore, regression may provide part of the explanation of the limited gains on the second testing. Conversely, for the lowest one-quarter, there is ample test ceiling and some increase in mean scores which is attributed to regression rather than to actual gains.

Performance by ability quarters by sex. Table XI completes the presentation of data for CEEB for the first semester, 1964-65. This table corresponds to Table VIII for COOP, both showing comparisons between sexes within subgroups at each of four ability levels. The right-hand portion of Table XI shows the superiority (or inferiority) of means for females as compared to means for males. At each of the four ability levels there are two comparisons: experimental males with experimental females and control males with control females. Of the eight resulting comparisons in September, the one that is noteworthy is within the experimental subgroup in the lowest quarter: the mean for females was 28.39 higher than the mean for males. In the three highest quarters the female-male disparity was not conspicuously different from zero in terms of the scale on which standard scores are reported and favored the males as often as the females.

An inspection of the male-female comparisons in January suggests two observations: one is that in the lowest one-quarter, the experimental females maintained the significant superiority which they displayed in September; the other is that in the highest one-quarter, the control females showed a substantial (52.10) mean superiority in contrast to a near-zero (3.86) superiority in September. The marked superiority of the highest-quarter control females in January resulted more from a substantial negative change (-34.93) on the part of the males than from the modest (13.30) positive change over the semester by the females.

TABLE XI

PERFORMANCE ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TEST BY MATCHED PAIRS BY SEX IN SEPTEMBER 1964 AND JANUARY 1965 AT EACH OF FOUR Z-SCORE[1] LEVELS: COMBINED INSTITUTIONS

Subgroup	Level	N	Sept. Z-Score Rating Mean	S.D.	Sept. CEEB Std. Rating Mean	S.D.	Jan. CEEB Std. Rating Mean	S.D.	CEEB Jan. minus Sept. Mean	r	t-Ratio	df	CEEB, Female minus Male September Diff.	t-Ratio	January Diff.	t-Ratio	df
Exp.Male	Highest 1/4	27	117.93	4.40	577.81	50.69	588.78	52.49	10.96	0.45	1.033	26	6.26	0.620	1.51	0.108	124
Exp.Female		99	122.73	7.06	584.07	44.76	590.29	66.62	6.22	0.32	0.913	98					
Cont.Male	"	27	117.74	3.90	581.56	44.02	546.63	64.82	-34.93	0.07	2.356*	26	3.86	0.393	52.10	3.452*	124
Cont.Female		99	122.53	6.93	585.42	45.27	598.73	70.04	13.30	0.26	1.803	98					
Exp.Male	2nd High 1/4	65	105.75	3.59	511.54	42.18	525.83	58.75	14.29	0.35	1.934	64	-6.73	1.034	1.70	0.177	176
Exp.Female		113	105.50	3.45	504.81	41.28	527.53	62.67	22.75	0.20	3.554*	112					
Cont.Male	"	65	105.66	3.86	507.25	37.91	521.83	66.58	14.58	0.22	1.695	64	-5.06	0.763	7.44	0.655	176
Cont.Female		113	105.61	3.60	502.19	44.64	529.27	75.75	27.07	0.37	3.957*	112					
Exp.Male	3rd High 1/4	65	93.88	3.74	444.83	44.75	483.57	66.46	38.74	0.12	4.110*	64	-5.10	0.674	-1.09	1.473	157
Exp.Female		94	93.80	3.50	439.73	47.85	482.48	55.73	42.74	0.27	6.571*	93					
Cont.Male	"	65	93.65	3.86	434.14	53.70	480.77	61.76	46.63	0.24	5.232*	64	11.06	0.111	-6.49	0.679	157
Cont.Female		94	93.93	3.38	445.20	40.33	474.28	56.89	29.07	0.16	4.361*	93					
Exp.Male	Lowest 1/4	78	77.24	8.60	370.59	53.17	420.60	56.67	50.01	0.13	6.059*	77	28.39	3.106*	26.40	2.663*	132
Exp.Female		56	81.86	4.72	398.98	49.82	447.00	55.48	48.02	0.09	5.006*	55					
Cont.Male	"	78	77.44	8.88	383.82	58.61	448.29	54.77	64.47	0.46	9.603*	77	13.02	1.352	2.44	0.248	132
Cont.Female		56	82.17	4.91	396.84	43.41	450.73	56.89	53.89	0.16	5.817*	55					

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

*Significant at 0.05 level (two-tailed test).

It is desirable to compare CEEB evidence in Tables IX, X, XI with the COOP evidence in Tables VI, VII, and VIII. These two tests were the ones which yielded September-to-January change scores (for the themes, it was not appropriate to compute change scores. See discussion page 42). Do students who have received a semester of instruction in freshman composition score higher as a group on these objective tests at the end of the semester than comparable students who have _not_ received such instruction? The answer is "Yes" if one judges by COOP scores, "No" if one judges by CEEB. Do females and males benefit about equally from such instruction? The answer is "Yes" for both the COOP and the CEEB. Do students at four ability levels benefit about equally from such instruction? The answer appears to be "No" for both the COOP and the CEEB.

Theme Rating, September 1964 and January 1965

Overall performance. Table XII presents theme performance for the 597 matched pairs of students who completed the tests through the first semester (September 1964-January 1965). Theme performance is generally considered to be the most direct measure of writing. Theme evidence for this study is unique also in that the matching procedures used required that the experimental and control means in September be identical. Interpretation of January means of experimentals and controls is therefore free of the qualifications which would have been necessary had there been unequal September means and standard deviations.

The N's involved in Table XII are all reasonably large. The two largest subgroups are the 597 experimental students and the 597 control students, that is, the 597 matched pairs. Data in the top portion of Table XII show that at the end of the first semester of college, the mean theme performance of the controls was 0.27 higher than that for the experimentals; the obtained t-value was 2.299, and significant. _Students receiving instruction such as that given in the first semester of freshman English composition performed significantly better on the theme than students who did not receive such instruction_. The variable under investigation was the presence or absence of instruction in freshman composition. In this experiment, such instruction had a positive influence on student writing performance.

It will be recalled that for the analyses on COOP and CEEB there was a significant advantage on the COOP for those who had received instruction, but not on the CEEB.

TABLE XII

THE PERFORMANCE OF 597 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

Subgroup	N	Degrees of Freedom	September Theme Total Mean	September Theme Total S. D.	Difference in September Means Control minus Experimental Diff.	r	t-Ratio	January Theme Total Mean	January Theme Total S. D.	Difference in January Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	597	596	9.15	2.14	0.00	1.00	0.000	9.69	2.47	0.27	0.31	2.299*
Controls	597		9.15	2.14				9.96	2.45			
Exp.Males	235	234	8.54	2.24	0.00	1.00	0.000	8.77	2.61	0.45	0.30	2.294*
Cont.Males	235		8.54	2.24				9.22	2.48			
Exp.Females	362	361	9.54	1.98	0.00	1.00	0.000	10.29	2.17	0.15	0.22	1.053
Cont.Females	362		9.54	1.98				10.44	2.30			

	September Diff. in Means Female minus Male Diff.	t-Ratio	January Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	1.00	5.690*	1.52	7.681*	596
Controls	1.00	5.690*	1.22	6.134*	596

*Significant at 0.05 level (two-tailed test).

Although this situation prevails for both males and females, the middle portion of Table XII reveals that the overall superiority in means for controls is due more to scores made by males than to those made by females. In January, control males had a mean total theme score which was 0.45 higher than that for the experimental males, whereas the control females had an observed superiority of 0.15 over the experimental females. The 0.45 for males was associated with a significant t-value. Thus the generalization above regarding the positive effect on college freshmen of one semester of instruction in composition may be refined by saying that these results are more distinctly characteristic of the male students than of the female students.

The bottom portion of Table XII compares the performance of males and females within the experimental subgroup and the control subgroup. These data relate to the question of whether, on this direct test of writing performance, the females perform better than the males. The mean difference in favor of the experimental females in September was 1.00. Similarly, the mean difference in favor of the control females was 1.00. These differences are significant.

How, then, do males and females compare in theme scores after a semester of college? The females have increased their superiority over the males, the difference in mean performance being 1.52 within the experimental subgroup and 1.22 within the control subgroup. Change scores on themes cannot be legitimately computed, owing to the difference in topic and time of evaluation; therefore one cannot, as was possible on the COOP and CEEB, analyze the change in score between the beginning and the end of the semester.

A number of comments are pertinent. In the fundamental comparison--597 experimental students against 597 control students on January theme--the obtained mean difference of 0.27 favoring the controls was significant. The 235 control males contributed more than did the 362 control females to the January finding of overall superiority of controls over experimentals. That is, control males surpassed experimental males (0.45) to a greater extent than did control females surpass experimental females (0.15).

The magnitude of these various between-subgroup differences in theme score means may be thought of in terms of mean-difference-necessary-for-significance of about 0.25 to 0.30. Thus one of the striking facts is the superiority of females over males, a condition which prevailed at the outset and is usually accentuated during the first semester of college attendance. There is a consistent tendency for the females to perform better than the males in group comparisons on all three criterion measures: COOP, CEEB, and Total Theme Rating.

The total sample of 597 matched pairs was composed of the following five institutional N's:

Institution	Male Pairs	Female Pairs	Total
1	65	135	200
2	65	62	122
3	34	90	124
4	9	11	20
5	62	64	126
Totals	235	362	597

It is helpful to examine the extent to which the findings for the combined institutions were observed consistently among the five constituent institutions. In four of the five institutions, at the end of the first semester, the control subgroup mean exceeded the experimental subgroup mean. However, in only one of these four institutions was the mean difference, in favor of the controls, statistically significant. In all five of the institutions, the females surpassed the males at the beginning of the freshman year, and the differential increased by the end of the semester.

Performance by ability quarters. Table XIII contains total theme means for the experimental students and the control students at each of four ability levels. Since the means in September were identical for any matched pair or subgroup of matched pairs, it is especially informative to look at the January means. It will be noticed that the control subgroup superiority of 0.27 for the total group (bottom portion of the table) resulted from the lowest quarter scores (mean difference of 0.67) and from the second high quarter scores (mean difference of 0.31). In the top quarter and the third high quarter the experimental and control students were essentially the same.

Performance by ability quarters by sex. From Table XIV it is possible to see the mean January theme ratings of males and females within the experimental and within the control subgroups at each of four ability levels. These ability levels were established on the basis of a combination of September COOP and CEEB scores. Table XIV also contains additional descriptive information: the Z-score means which were used in establishing the four ability levels, and the means and differences in means on September theme ratings.

The three columns at the right-hand side of Table XIV present the key comparisons. Generally in the eight comparisons, the mean of January theme ratings for the females was higher than for the males. The single exception was with the experimental students in the highest quarter, where the means for males was 0.08 higher than the mean for females. Another way of summarizing the

TABLE XIII

THE PERFORMANCE OF 597 MATCHED PAIRS OF STUDENTS ON TOTAL THEME IN SEPTEMBER 1964 AND JANUARY 1965 AT EACH OF FOUR Z-SCORE[1] LEVELS; MEANS, STANDARD DEVIATIONS, DIFFERENCES IN MEANS, AND t-RATIOS: COMBINED INSTITUTIONS

			September		January		Difference in Jan. Means Control minus Experimental			
Quarter	Subgroup	N	Mean	S.D.	Mean	S.D.	Difference	r	t-Ratio	d.f.
Highest 1/4	Experimental	126	10.44	1.85	11.20	2.21	-0.01	0.17	0.031	125
"	Control	126	10.44	1.85	11.21	2.30				
Second High 1/4	Experimental	178	9.61	1.90	10.19	2.10	0.31	0.05	1.387	177
"	Control	178	9.61	1.90	10.50	2.27				
Third High 1/4	Experimental	159	8.62	1.92	9.25	2.34	0.09	0.27	0.422	158
"	Control	159	8.62	1.92	9.34	2.33				
Lowest 1/4	Experimental	134	7.95	2.11	8.14	2.25	0.67	0.26	2.862*	133
"	Control	134	7.95	2.11	8.81	2.18				
Total Group	Experimental	597	9.15	2.14	9.69	2.47	0.27	0.31	2.299*	596
"	Control	597	9.15	2.14	9.96	2.45				

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

*Significant at 0.05 level (two-tailed test).

TABLE XIV

PERFORMANCE ON THEME BY MATCHED PAIRS BY SEX IN SEPTEMBER 1964 AND JANUARY 1965 AT EACH OF FOUR Z-SCORE[1] LEVELS: COMBINED INSTITUTIONS

			Sept. Z-Score Rating		Sept. Theme Rating		Jan. Theme Rating		Theme, Female minus Male				
									September		January		
Subgroup	Level	N	Mean	S.D.	Mean	S.D.	Mean	S.D.	Diff.	t-Ratio	Diff.	t-Ratio	df
Exp.Male	Highest	27	117.93	4.40	9.96	1.99	11.26	2.07	0.62	1.524	-0.08	0.160	124
Exp.Female	1/4	99	122.73	7.06	10.58	1.79	11.18	2.25					
Cont.Male	"	27	117.74	3.90	9.96	1.99	10.48	2.70	0.62	1.524	0.92	1.861	124
Cont.Female		99	122.53	6.93	10.58	1.79	11.40	2.13					
Exp.Male	Second	65	105.75	3.59	9.51	2.08	9.86	2.37	0.15	0.525	0.51	1.563	176
Exp.Female	High 1/4	113	105.50	3.45	9.66	1.78	10.37	1.91					
Cont.Male	"	65	105.66	3.86	9.51	2.08	10.08	2.32	0.15	0.525	0.66	1.891	176
Cont.Female		113	105.61	3.60	9.66	1.78	10.74	2.21					
Exp.Male	Third	65	93.88	3.74	8.25	1.86	8.29	2.47	0.63	2.076*	1.61	4.513*	157
Exp.Female	High 1/4	94	93.80	3.50	8.88	1.91	9.90	2.00					
Cont.Male	"	65	93.65	3.86	8.25	1.86	8.72	2.28	0.63	2.076*	1.05	2.829*	157
Cont.Female		94	93.93	3.38	8.88	1.91	9.77	2.26					
Exp.Male	Lowest	78	77.24	8.60	7.50	2.13	7.40	2.02	1.07	2.967*	1.78	4.864*	132
Exp.Female	1/4	56	81.86	4.72	8.57	1.92	9.18	2.15					
Cont.Male	"	78	77.44	8.88	7.50	2.13	8.49	2.30	1.07	2.967*	0.78	2.059*	132
Cont.Female		56	82.17	4.91	8.57	1.92	9.27	1.91					

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

*Significant at 0.05 level (two-tailed test).

facts portrayed in Table XIV is by noting that the superiority of females over males was most prominent in the two lowest quarters--and within those two quarters, more so within the experimental subgroup than within the control subgroup.

The following tabulation shows, for the 597 matched pairs who completed the first semester, the number and percentage of males and females who had been in the upper half and in the lower half on the September ability distribution according to two determiners of ability levels: the Z-score, a combination of two objective tests; and Total Theme Rating.

PROPORTION OF MALES AND FEMALES IN JANUARY WHO WERE IN UPPER HALF AND LOWER HALF OF SAMPLE IN SEPTEMBER

	Number in January	Men		Women	
		Number	Percent	Number	Percent
On September Z-score:					
Upper Half	304	92	30	212	70
Lower Half	293	143	49	150	51
Upper + Lower	597	235	39	362	61
On September Theme Rating:					
Upper Half	359	117	33	242	67
Lower Half	237	117	49	120	51
Upper + Lower	596	234	39	362	61

The analyses by ability level thus far have all been based on the Z-score (objective-test based) ability levels. For the 597 matched pairs who completed the first semester testing, 235 (39.4%) were males. Among the original set of 1,040 matched pairs in September the percentage of males was 40.6. It is evident that about the same proportion of males and females persisted through the first semester. It is noted that of the 235 persisting men, 39.1 percent were in the upper half of the objective-test distribution, while 60.9 percent were in the lower half. For the females, the corresponding percentages were 58.6 in the upper half and 41.4 in the lower half.

Persistence of matched pairs of students of the upper one-half in ability and the lower one-half in ability was dependent upon the ability measure employed. For the Z-score measure of ability, persistence was approximately equal for the upper and lower halves (304, upper, 293, lower). For the essay measure of ability, persistence was greater for the upper half (359) than for the lower half (237).

If the persisting men are categorized in terms of the September theme distribution, it is noted that 50 percent were in the upper half and 50 percent in the lower half (117 in each); for the persisting women the percentages were 66.9 (242) and 33.1 (120). Thus among the 597 matched pairs who persisted through the first semester, the superiority of the female portion over the male portion in terms of September performance was slightly greater for the theme criterion than for the objective-test criterion.

Table XV contains information concerning the 596 matched pairs of students whose performances were portrayed in the tabulation on page 74. From the data of Table XV it is possible to see the make-up of the sample by ability quarters as determined by September theme rating. Within each quarter there is an indication of the number of males and females and the mean scores within each subgroup on September theme and on Z-score (two September objective tests combined). Since the ability quarters were established in terms of theme ratings (in contrast to Z-score on preceding tables) and since matching between subgroups was perfect on sex and theme rating, the differences among ability levels are sharper in the "Theme" column than in the "Z-score" column. The mean differences between adjacent quarters on theme for males were 2.12, 1.96, and 2.19; and for females 2.26, 1.87, and 1.96.

First Year Sample

Table XV concluded a series of tables which presented the performance of the 597 pairs of students who completed the testing at the end of the first semester (September 1964-January 1965). The next series of tables, beginning with Table XVI and extending through Table XXXIII, will present data for the 365 matched pairs who completed the entire 1964-65 academic year, and were tested in May 1965. These students constitute a subgroup of the 597 matched pairs whose performance has just been summarized. These 365 matched pairs of students were tested in May as well as in September and January.

COOP, First Semester

Overall performance. Table XVI presents COOP performance of the 365 matched pairs at the beginning and at the end of the first semester, 1964-65. Four main comparisons are given: between subgroups in September; within subgroups, September to January;

TABLE XV

THE 596 MATCHED PAIRS WHO PERSISTED THROUGH JANUARY 1965: MEANS ON SEPTEMBER THEME AND SEPTEMBER Z-SCORE[1] BY SEX BY SUBGROUP BY ABILITY LEVEL ESTABLISHED ON SEPTEMBER THEME RATING

Quarter	N	Sex	Subgroup	Mean, Sept. Theme	Mean, Sept. Z-Score
Highest 1/4	43	Male	Experimental	11.78	104.81
"	43	"	Control	11.78	104.46
"	114	Female	Experimental	11.84	109.61
"	114	"	Control	11.84	109.69
2nd High 1/4	74	Male	Experimental	9.55	98.12
"	74	"	Control	9.55	98.12
"	128	Female	Experimental	9.58	104.48
"	128	"	Control	9.58	104.66
3rd High 1/4	69	Male	Experimental	7.59	90.75
"	69	"	Control	7.59	90.91
"	96	Female	Experimental	7.71	98.11
"	96	"	Control	7.71	98.08
Lowest 1/4	48	Male	Experimental	5.40	84.00
"	48	"	Control	5.40	83.83
"	24	Female	Experimental	5.75	91.00
"	24	"	Control	5.75	90.75
Total Group	234	Male	Experimental	8.54	94.40
"	234	"	Control	8.54	94.36
"	362	Female	Experimental	9.54	103.51
"	362	"	Control	9.54	103.58

---

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

TABLE XVI

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

Subgroup	N	Cooperative Eng. Test Converted Score September Mean	S. D.	January Mean	S. D.	Difference in Means, Jan. minus September	r	t-Ratio	Difference in Sept. Means, Control minus Experimental Diff.	r	t	Difference in Jan. Means, Control minus Experimental Diff.	r	t	Degrees of Freedom
Experimental	365	162.60	7.30	164.51	7.45	1.91	0.72	6.561*	0.19	0.80	0.791	1.25	0.61	3.733*	364
Control	365	162.79	7.40	165.76	7.04	2.97	0.69	9.949*							
Exp.Males	134	160.51	7.35	162.77	7.72	2.26	0.69	4.394*	0.13	0.77	0.291	0.54	0.64	1.000	133
Cont.Males	134	160.63	7.60	163.31	7.05	2.68	0.56	4.487*							
Exp.Females	231	163.81	6.99	165.52	7.09	1.71	0.71	4.873*	0.23	0.80	0.795	1.66	0.56	3.921*	230
Cont.Females	231	164.04	6.97	167.19	6.64	3.14	0.74	9.779*							

	September Diff. in Means Female minus Male Diff.	t-Ratio	January Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	3.30	4.262*	2.75	3.453*	363
Controls	3.41	4.342*	3.88	5.236*	363

*Significant at 0.05 level (two-tailed test).

between subgroups in January; between the male and female portions of these subgroups. Table XVI may be compared with Table VI, which presents COOP data for the 597 matched pairs who completed the January 1965 testing. In both tables the means for January show a superiority for the control subgroup over the experimental subgroup. For the 597, the superiority was 1.01, and for the 365, the superiority was 1.25. A noticeable variation between males and females in the two samples was that, for the 365 matched pairs in January (Table XVI), the control males were not substantially superior to the experimental males, whereas for the 597 matched pairs (Table VI), the control males were significantly superior to the experimental males in January.

One of the functions of this comparison of partially overlapping samples is to permit inferences regarding the nature of the matched pairs who were lost from the experiment between January 1965 and May 1965. Except for the control males, who constituted only about 37 percent of the control subgroup in Table XVI, the two samples, 365 pairs and 597 pairs, are very similar. Consequently, the 232 pairs who vanished over the second semester apparently were not appreciably different from the 365 matched pairs who remained. To put it differently, the loss of the 232 matched pairs apparently did not alter the representativeness of the remaining sample.

Performance by ability quarters. Table XVII, as compared to Table XVI, shows the experimental-control comparisons at each of four ability levels. A picture of the gains on COOP during the first semester is presented in the column headed "Difference in Means, January minus September." As was the case for the parent sample of 597 matched pairs (Table VII, page 59), the semester gains among the sample of 365 matched pairs was relatively large in the lowest quarter of ability. The 80 experimental students in the lowest one-quarter showed a mean gain of 3.80; the 80 control students in the same quarter had a mean gain of 5.34. These changes may be compared to 1.91 and 2.97 for the total group of 365 matched pairs.

For the overall sample of 365 pairs, the January COOP means showed a superiority of 1.25 for the control subgroup. This overall mean difference resulted from the following mean differences at the four ability levels: 0.46 (highest), 2.16, 0.31, 1.75 (lowest). The differences for the second and the fourth quarters are significant. However, since there is no definite trend, the mean differences by ability level should be interpreted cautiously.

TABLE XVII

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965, BY Z-SCORE[1] QUARTERS; MEANS, STANDARD DEVIATIONS, AND t-RATIOS: COMBINED INSTITUTIONS

			Sept. Z-Score		COOP Eng. Conv. Score Sept. 1964		COOP Eng. Conv. Score Jan. 1965		Difference in Means, January 1965 minus Sept. 1964			Diff. in Sept. Means, Control minus Exp.			Diff. in Jan. Means, Control minus Exp.			
Quarter	Subgroup	N	Mean	S. D.	Mean	S. D.	Mean	S. D.	Diff.	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	df
Highest 1/4	Exper.	81	121.83	6.79	171.44	4.12	172.47	5.57	1.02	0.46	1.762	0.01	0.43	0.025	0.46	0.31	0.672	80
"	Cont.	81	121.59	6.75	171.43	4.07	172.93	4.67	1.49	0.58	3.312*							
2nd High 1/4	Exper.	116	105.93	3.51	164.47	3.35	165.49	5.55	1.03	0.30	1.983*	0.55	0.31	1.396	2.16	0.13	3.261*	115
	Cont.	116	106.01	3.62	165.02	3.85	167.65	5.16	2.63	0.28	5.125*							
3rd High 1/4	Exper.	90	94.11	3.66	160.41	3.90	162.53	4.26	2.12	0.35	4.280*	-0.12	0.21	0.232	0.31	0.25	0.556	89
	Cont.	90	94.10	3.77	160.29	4.00	162.84	3.39	2.56	0.13	4.349*							
Lowest 1/4	Exper.	80	80.35	6.74	153.55	4.70	157.35	5.98	3.80	0.33	5.384*	0.19	0.45	0.338	1.73	0.38	2.345*	79
"	Cont.	80	80.84	6.87	153.74	4.69	159.08	5.72	5.34	0.17	7.024*							
Total Group	Exper.	365	100.88	15.34	162.60	7.30	164.51	7.45	1.91	0.72	6.561*	0.19	0.80	0.791	1.25	0.61	3.733*	364
"	Cont.	365	100.96	15.16	162.79	7.40	165.76	7.04	2.97	0.69	9.949*							

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

*Significant at 0.05 level (two-tailed test).

CEEB, First Semester

Overall performance. Table XVIII presents the same comparisons of CEEB data for the first semester that Table XVI gave of COOP data for the 365 matched pairs surviving in May 1965.

For the experimental subgroup and the control subgroup the mean gains on CEEB, September to January, are of the order of 30 standard rating points--30.14 for the experimentals and 28.72 for the controls. These gains are significant. Similarly, the analysis by sex shows significant mean gains for males and for females. An inspection of the January minus September column of Table XVIII shows that the control males had the smallest mean gain, 20.52.

The key comparison is between experimental mean and control mean subgroups in January. The two means were very close in January, as they had been in September: the January difference of -2.79 is not significant. The comparison of subgroups by sex also fails to reveal any significant mean differences. The correlations of CEEB scores within subgroups and between subgroups are in line with expectation. The September-January comparisons yield related r's of the order of 0.60. The between-subgroup r's, also related r's, were approximately 0.80 in September and 0.40 in January.

The bottom portion of Table XVIII compares CEEB means for males and females within the experimental subgroup and within the control subgroup. The superiority of the females was about 26 standard rating points in September and increased to 31.40 within the experimental subgroup and 38.96 within the control subgroup in January, both differences being significant.

There was substantial agreement in January between the findings for the 365 matched pairs available for the first two semesters (Table XVIII) and the 597 matched pairs available at the end of the first semester, Table IX, page 63).

Performance by ability quarters. From Table XIX, containing CEEB data for the first semester by ability levels, it is evident that mean gains were largest in the two lowest quarters. For the larger sample (597--Table X, page 65) the semester mean gain was on the order of 30 standard rating points. Table XIX shows that for the 365 pairs, the mean gains were 20 standard rating points or less for students in the highest two quarters in ability, whereas for the students in the lower two quarters of ability, the four mean gains were 49.77 (experimental), 33.80 (control); and 40.72 (experimental), 62.10 (control).

TABLE XVIII

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

		College Entrance Examination Board Standard Rating				Difference in Means Jan. minus			Diff.in September Means, Control minus Experimental			Diff.in January Means, Control minus Experimental			
		September		January											
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	365	484.19	83.25	514.33	81.89	30.14	0.67	8.617*	-1.38	0.81	0.513	-2.79	0.47	0.628	364
Control	365	482.81	82.85	511.54	82.43	28.72	0.59	7.318*							
Exp.Males	134	467.49	85.47	494.46	82.47	26.97	0.73	5.033*	-1.13	0.78	0.230	-7.58	0.36	0.986	133
Cont.Males	134	466.36	86.65	486.88	74.00	20.52	0.54	3.036*							
Exp.Females	231	493.87	80.36	525.86	79.30	31.98	0.62	6.992*	-1.52	0.82	0.483	-0.02	0.49	0.003	230
Cont.Females	231	492.36	79.01	525.84	83.68	33.48	0.60	6.991*							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	26.38	2.945*	31.40	3.583*	363
Controls	26.00	2.916*	38.96	4.458*	363

*Significant at 0.05 level (two-tailed test).

TABLE XIX

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965, BY Z-SCORE[1] QUARTERS: MEANS, STANDARD DEVIATIONS, AND t-RATIOS: COMBINED INSTITUTIONS

Quarter	Subgroup	N	Sept. Z-Score Mean	Sept. Z-Score S. D.	CEEB Stan. Rating Sept. 1964 Mean	CEEB Stan. Rating Sept. 1964 S.D.	CEEB Stan. Rating Jan. 1965 Mean	CEEB Stan. Rating Jan. 1965 S.D.	Difference in Means, Jan. 1965 minus Sept. 1964 Diff.	Difference in Means, Jan. 1965 minus Sept. 1964 r	Difference in Means, Jan. 1965 minus Sept. 1964 t-Ratio	Diff. in Sept. Means, Control minus Exp. Diff.	Diff. in Sept. Means, Control minus Exp. r	Diff. in Sept. Means, Control minus Exp. t-Ratio	Diff. in Jan. Means, Control minus Exp. Diff.	Diff. in Jan. Means, Control minus Exp. r	Diff. in Jan. Means, Control minus Exp. t-Ratio	df
Highest 1/4	Exper.	81	121.83	6.79	584.32	46.35	595.96	59.40	11.64	0.27	1.610	-1.57	0.41	0.280	-9.40	0.20	1.008	80
"	Cont.	81	121.59	6.75	582.75	45.77	586.57	71.40	3.81	0.19	0.442							
2nd High 1/4	Exper.	116	105.93	3.51	511.94	41.18	532.23	64.33	20.29	0.28	3.305*	5.97	0.43	1.404	-7.91	0.09	0.893	115
	Cont.	116	106.01	3.62	505.97	44.12	524.32	76.35	18.35	0.30	2.587*							
3rd High 1/4	Exper.	90	94.11	3.66	441.67	46.41	491.43	52.82	49.77	0.32	8.086*	1.36	0.33	0.226	-14.61	0.02	1.781	89
	Cont.	90	94.10	3.77	443.02	51.35	476.82	57.72	33.80	0.23	4.685*							
Lowest 1/4	Exper.	80	80.35	6.74	391.70	52.74	432.43	58.47	40.72	0.19	5.114*	2.59	0.53	0.438	23.96	0.24	2.981*	79
"	Cont.	80	80.84	6.87	394.29	55.00	456.39	57.54	62.10	0.35	8.611*							
Total Group	Exper.	365	100.88	15.34	484.19	83.25	514.33	81.89	30.14	0.67	8.617*	-1.38	0.81	0.513	-2.79	0.47	0.628	364
"	Cont.	365	100.96	15.16	482.81	82.85	511.54	82.43	28.72	0.59	7.318*							

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

*Significant at 0.05 level (two-tailed test).

The comparison between the 365 experimentals and the 365 controls at the end of the first semester shows two interesting facts about the lowest one-quarter: it was the only one-quarter in which the control mean excelled the experimental mean, and this was the only one of the four quarters at which the between-subgroups difference is significant.

In summary, these CEEB data for the first semester of the 365 matched pairs indicate overall similarity between the performance of the experimental and the control subgroups. On end-of-semester comparisons between subgroups, the control subgroup excelled significantly in the lowest quarter, while the experimental subgroup had a slight but not significant advantage in the upper three quarters. The two lowest quarters showed greater gain than the two upper quarters.

Theme Rating, First Semester

Overall performance. The data in Table XX, covering the theme ratings for the first semester for the 365 matched pairs surviving in May 1965, may be related to the first semester data for the 597 matched pairs surviving in January 1965 (Table XII, page 69). This is helpful in considering the extent to which the matched pairs finishing the full freshman year are representative of the matched pairs completing only the first semester of the freshman year. The first two lines in Table XX show data for 365 matched pairs who were included among the 597 matched pairs depicted in the first two lines of Table XII. It may be noted in Table XX that in September the mean theme rating for the smaller group (N=365) was 9.23, whereas the corresponding figure for the larger subgroup (N=597) was 9.15. The persistence rate for males was slightly less than for females: 57.0 percent and 63.8 percent, respectively. Such evidence of a slight selective factor in this kind of longitudinal study would be expected.

It was noted, from Table XII, that at the end of the first semester the 597 control students had a mean theme rating which was 0.27 higher than that of the mean for the 597 experimental students, and that this difference was significant. In Table XX it is shown that for the 365 matched pairs, the mean difference on January theme in favor of the controls was 0.26. Coupled with a relatively large standard error, this mean difference is not quite significant.

In general, however, the analysis by sex for the 365 matched pairs (Table XX) yielded results which were consistent with those for the 597 matched pairs (Table XII). In both tables, the September means for the females were higher than those for the males,

TABLE XX

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

			September Theme Total		Difference in September Means Control minus Experimental			January Theme Total		Difference in January Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Mean	S. D.	Diff.	r	t-Ratio	Mean	S. D.	Diff.	r	t-Ratio
Experimentals	365	364	9.23	2.01	0.00	1.00	0.000	9.95	2.38	0.26	0.26	1.730
Controls	365		9.23	2.01				10.21	2.35			
Exp.Males	134	133	8.83	2.14	0.00	1.00	0.000	9.11	2.55	0.46	0.27	1.790
Cont.Males	134		8.83	2.14				9.57	2.36			
Exp.Females	231	230	9.46	1.90	0.00	1.00	0.000	10.44	2.13	0.14	0.19	0.774
Cont.Females	231		9.46	1.90				10.58	2.26			

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.63	2.909*	1.33	5.329*	363
Controls	0.63	2.909*	1.01	4.026*	363

*Significant at 0.05 level (two-tailed test).

and in both cases, the January means showed an even greater superiority for the females. That is, the initial difference did not disappear, as it is sometimes averred to do.

Performance by ability quarters. Data in Table XXI show how the overall mean difference of 0.26 on total theme rating in January (10.21, control; 9.95, experimental) was apportioned among the four ability levels. The overall mean difference is not quite significant, but in the lowest quarter the observed mean difference of 0.69 is significant. However, the fact that the observed mean difference for the second quarter approached significance suggests the complexity of control-experimental comparisons by ability level.

Two instructive facts are included among the auxiliary evidence in Table XXI. One of these facts is the correlations between the total theme ratings of September and January, within subgroups within ability levels, and overall. Within each of the ability levels, the r's are low, typically in the 0.20's and 0.30's. The overall r's are 0.44 and 0.34. The relative smallness of all of the r's may be due in part to the relative unreliability of theme ratings (see discussion of reliability, page 43). The second fact concerns the influence of range of talent on the magnitude of correlation coefficients; the r's within the ability levels tend to be lower than the overall r's.

COOP, Second Semester

Overall performance. Table XXII is the first in a series of tables showing data for the second semester of the freshman year. The data in this series of tables involve the same 365 matched pairs dealt with in Tables XVI through XXI.

From the column "Difference in Means, May minus January" (Table XXII), it is seen that gains on COOP during the semester were between 1-1/2 and 2 converted score points, and that these gains are all significant. Gains during the semester were roughly similar for experimentals and controls, and for males and females.

The control subgroup began the second semester with a significant mean superiority over the experimentals (1.25) and ended the semester with a smaller, but still significant, mean difference over the experimentals (0.79). The breakdown of the data by sex shows that, among males, the controls began the semester 0.54 higher than the experimentals and ended the semester 0.28 higher; on neither of the two occasions is the difference significant. For the females, the controls had a significantly higher mean than the experimentals at the beginning (1.66) and at the end (1.10).

TABLE XXI

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON TOTAL THEME IN SEPTEMBER 1964 AND JANUARY 1965 AT EACH OF FOUR Z-SCORE[1] LEVELS; MEANS, STANDARD DEVIATIONS, DIFFERENCES IN MEANS, AND t-RATIOS: COMBINED INSTITUTIONS

			Sept.1964 Z-Score		Sept.1964 Theme		Jan.1965 Theme		Difference in Jan. Means Control minus Experimental			
Quarter	Subgroup	N	Mean	S.D.	Mean	S.D.	Mean	S.D.	Difference	r	t-Ratio	d.f.
Highest 1/4	Experimental	81	121.83	6.79	10.35	1.79	11.49	2.03	-0.11	0.07	0.341	80
"	Control	81	121.59	6.75	10.35	1.79	11.38	2.23				
Second High 1/4	Experimental	116	105.93	3.51	9.47	1.89	10.16	2.18	0.47	0.08	1.694	115
"	Control	116	106.01	3.62	9.47	1.89	10.63	2.24				
Third High 1/4	Experimental	90	94.11	3.66	8.71	1.80	9.63	2.15	-0.07	0.17	0.225	89
"	Control	90	94.10	3.77	8.71	1.80	9.57	2.20				
Lowest 1/4	Experimental	80	80.35	6.74	8.24	2.02	8.45	2.19	0.69	0.27	2.371*	79
"	Control	80	80.84	6.87	8.24	2.02	9.14	2.08				
Total Group	Experimental	365	100.88	15.34	9.23	2.01	9.95	2.38	0.26	0.26	1.730	364
"	Control	365	100.96	15.16	9.23	2.01	10.21	2.35				

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

*Significant at 0.05 level (two-tailed test).

TABLE XXII

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

Subgroup	N	Cooperative Eng. Test Converted Score January Mean	S. D.	May Mean	S. D.	Difference in Means, May minus January	r	t-Ratio	Difference in Jan. Means, Control minus Experimental Diff.	r	t	Difference in May Means, Control minus Experimental Diff.	r	t	Degrees of Freedom
Experimental	365	164.51	7.45	166.50	6.82	1.99	0.69	6.675*	1.25	0.61	3.733*	0.79	0.55	2.283*	364
Control	365	165.76	7.04	167.29	7.09	1.53	0.71	5.397*							
Exp.Males	134	162.77	7.72	164.63	7.22	1.87	0.68	3.583*	0.54	0.64	1.000	0.28	0.44	0.431	133
Cont.Males	134	163.31	7.05	164.91	6.68	1.60	0.64	3.152*							
Exp.Females	231	165.52	7.09	167.58	6.33	2.06	0.67	5.693*	1.66	0.56	3.921*	1.10	0.58	2.704*	230
Cont.Females	231	167.19	6.64	168.68	6.96	1.49	0.72	4.401*							

	January Diff. in Means Female minus Male Diff.	t-Ratio	May Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	2.75	3.453*	2.95	4.055*	363
Controls	3.88	5.236*	3.77	5.042*	363

*Significant at 0.05 level (two-tailed test).

The recurring fact is that the greatest differences in means on COOP are found in comparisons between females and males, regardless of whether they are experimentals or controls. Within the experimental subgroup, the superiority of the mean for females over the mean for males was 2.75 in January and 2.95 in May. The corresponding figures within the control subgroup were 3.88 and 3.77. All four of these differences in favor of the females are significant.

Performance by ability quarters. Data in Table XXIII enable the reader to identify any student ability levels at which the January-to-May gains on COOP are noteworthy. The lowest quarter of ability showed a relatively high mean gain for the experimentals (3.21) and a relatively low mean gain for the controls (1.09). At the beginning of the second semester, the mean of the lowest one-quarter of the control subgroup was 1.73 higher than the corresponding experimental subgroup mean. At the end of the second semester, the observed difference between these two lowest-quarter subgroups was considerably less, 0.40.

The second highest quarter of ability is unique in that both the January and the May between-subgroup differences favoring the controls are significant (2.16 and 1.23 respectively). The mean difference in May was 0.93 smaller than the mean difference in January.

It was only in the third high quarter that the control subgroup superiority was greater in May (1.41) than it had been in January (0.31); the 1.41 mean difference is not significant.

Thus, in general, the second semester of instruction did not maintain the superiority of the controls over the experimentals. During the second semester, the experimentals out-gained the controls, who continued to receive instruction.

In the analysis by ability level, the investigators were on the alert for patterns of performance which would have implications for curriculum and instruction. Thus if a noteworthy fact emerged for the top one-half or the bottom one-half, or the top quarter or the bottom quarter, one might see a possible application of such findings to exemption or sectioning. With COOP data for the second semester, no clear-cut pattern emerged. The lower half yielded two facts which illustrate the absence of a pattern: the relatively large gains by the controls of the third highest one-quarter and the experimentals of the lowest one-quarter. The investigators wonder whether such facts as those for this sample are chance findings or whether they are indicative of the actual situation for the population sampled.

TABLE XXIII

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN JANUARY 1965 AND MAY 1965, BY Z-SCORE[1] QUARTERS; MEANS, STANDARD DEVIATIONS, AND t-RATIOS: COMBINED INSTITUTIONS

			Sept. Z-Score		COOP Eng. Conv. Score Jan. 1965		COOP Eng. Conv. Score May 1965		Difference in Means, May 1965 minus Jan. 1965			Diff. in Jan. Means, Control minus Exp.			Diff. in May Means, Control minus Exp.			
Quarter	Subgroup	N	Mean	S. D.	Mean	S. D.	Mean	S. D.	Diff.	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	df
Highest 1/4	Exper.	81	121.83	6.79	172.47	5.57	173.84	5.09	1.37	0.49	2.266*	0.46	0.31	0.672	0.59	0.33	0.908	80
"	Cont.	81	121.59	6.75	172.93	4.67	174.43	5.02	1.51	0.65	3.307*							
2nd High 1/4	Exper.	116	105.93	3.51	165.49	5.55	167.55	4.75	2.06	0.50	4.240*	2.16	0.13	3.261*	1.23	0.23	2.137*	115
	Cont.	116	106.01	3.62	167.65	5.16	168.78	5.22	1.14	0.36	2.072*							
3rd High 1/4	Exper.	90	94.11	3.66	162.53	4.26	163.88	4.54	1.34	0.27	2.391*	0.31	0.25	0.556	1.41	-0.10	1.933	89
	Cont.	90	94.10	3.77	162.84	4.39	165.29	4.75	2.44	0.27	4.179*							
Lowest 1/4	Exper.	80	80.35	6.74	157.35	5.98	160.56	5.58	3.21	0.34	4.301*	1.73	0.38	2.345*	0.40	0.07	0.476	79
"	Cont.	80	80.84	6.87	159.08	5.72	160.16	5.39	1.09	0.50	1.744							
Total Group	Exper.	365	100.88	15.34	164.51	7.45	166.50	6.82	1.99	0.69	6.675*	1.25	0.61	3.733*	0.79	0.55	2.283*	364
"	Cont.	365	100.96	15.16	165.76	7.04	167.29	7.09	1.53	0.71	5.397*							

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

*Significant at 0.05 level (two-tailed test).

CEEB, Second Semester

Overall performance. Table XXIV contains second-semester CEEB data of the 365 matched pairs surviving in May 1965. Comparable CEEB data of these same pairs for the first semester are in Table XVIII, page 81 The within-semester analysis shows that during the first semester the mean gains by experimentals (30.14) and by controls (28.72) were very similar. During the second semester (Table XXIV) two noteworthy facts emerge: the gains were smaller than for the first semester and the controls outperformed the experimentals 12.74 to 2.65. Thus, for this sample, the increment for one year occurred primarily in the first semester--for the experimentals almost all of it and for the controls about two-thirds of it. All of this is reflected in the column showing t-ratio. During the first semester (Table XVIII) both the experimental subgroup and the control subgroup made highly significant gains, whereas during the second semester (Table XXIV) only the controls gained significantly.

For the 365 pairs of students, end-of-semester differences between means of the experimental subgroup and the control subgroup were not significant. At the end of the first semester the subgroup means differed by 2.79 favoring the experimentals; at the end of the second semester, they differed by 7.30 favoring the controls. The obtained mean difference of 7.30 standard rating points is associated with a t-value of 1.640, short of significance at the 0.05 level. For 364 degrees of freedom a t of 1.97 is required for significance. Under the given conditions of variability and correlation, an observed mean difference of 8.86 would be a significant difference. For example, had the control subgroup mean been 525.80 instead of 524.28, the difference would have attained significance.

What would be required to have a mean difference of 8.86 standard rating points? The CEEB has 100 to 110 items, depending on the form. An increase of one raw score point is typically associated with an increase of about six standard rating points. Thus if, on the average, members of one of these two subgroups had made one or two more correct responses than did their counterparts in the other subgroup, the resulting subgroup means would have differed significantly.

A reference distribution of student scores may serve as a vehicle for additional thinking about the within-subgroups and between-subgroups analyses presented in Tables IX, XVIII, and XXIV. The following tabulation shows, for each of six mean standard ratings, two for each testing period, the percentile rank which each such rating had in a distribution of September CEEB scores (N=4,159) and in a distribution of May CEEB scores (N=730).

TABLE XXIV

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in January Means, Control minus Experimental			Diff. in May Means, Control minus Experimental			
		January		May		May minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	January	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	365	514.33	81.89	516.98	82.01	2.65	0.65	0.738	-2.79	0.47	0.628	7.30	0.44	1.640	364
Control	365	511.54	82.43	524.28	78.94	12.74	0.64	3.537*							
Exp.Males	134	494.46	82.47	501.47	77.43	7.01	0.72	1.353	-7.58	0.36	0.986	-1.19	0.47	0.172	133
Cont.Males	134	486.88	74.00	500.28	77.51	13.40	0.57	2.187*							
Exp.Females	231	525.86	79.30	525.98	83.24	0.13	0.60	0.026	0.02	0.49	0.003	12.22	0.41	2.127*	230
Cont.Females	231	525.84	83.68	538.20	76.37	12.36	0.65	2.774*							

	January Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of
	Diff.	t-Ratio	Diff.	t-Ratio	Freedom
Experimentals	31.40	3.583*	24.51	2.774*	363
Controls	38.96	4.458*	37.92	4.536*	363

*Significant at 0.05 level (two-tailed test).

Subgroup	Rounded Mean Standard Rating	Percentile Rank in a Distribution of Student Scores in September 1964	and May 1965
Experimentals (Sept. 1964)	485	54+	38
Controls "	484	54	37
Experimentals (Jan. 1965)	514	65	50
Controls "	511	64	49
Experimentals (May 1965)	517	67	51
Controls	524	68	55

This kind of analysis also emphasizes the similarity of the observed means of experimentals and controls. In a distribution of student scores, the corresponding percentile rank differences were typically 1. For example, the January means of 514 and 511 correspond to September percentile ranks (among individual scores) of 65 and 64, respectively. The mean gains during the semester may be examined in a similar manner. If the typical end-of-semester score is placed in a beginning-of-semester distribution of student scores, it appears that the improvement is 10 or 11 percentile rank points. This finding--only a modest increment in performance as a result of an additional semester or year of instruction plus maturation--is illustrative of a rather general condition which is perhaps not sufficiently appreciated by teachers of English.

The present investigators had previously acknowledged this aspect of year-by-year instruction in curricular areas common to consecutive school levels.[8,9]

Performance by ability quarters. One function of Table XXV is to look beyond mean second semester gains of 2.65 (non-significant) standard rating points by the experimental subgroup and 12.74 (significant) by the control subgroup through examining second semester CEEB data by ability levels. At none of the four levels

---

[8] Jewell, Ross M., et al., Final Report of the Communication Experiment Conducted by the Department of Languages, Speech, and Literature of the Iowa State Teachers College, 1955-58. (Cedar Falls, Iowa, 1960).

[9] Jewell, Ross M., and Gordon J. Rhum, The Relative Effectiveness of Two Methods of Instruction in College Freshman Composition: Closed-Circuit Television and "Normal" Classroom. (Cedar Falls, Iowa: State College of Iowa, 1966).

TABLE XXV

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN JANUARY 1965 AND MAY 1965, BY Z-SCORE[1] QUARTERS: MEANS, STANDARD DEVIATIONS, AND t-RATIOS: COMBINED INSTITUTIONS

Quarter	Subgroup	N	Sept. Mean	Z-Score S. D.	CEEB Stan. Rating Jan. 1965 Mean	S.D.	CEEB Stan. Rating May 1965 Mean	S.D.	Difference in Means, May 1965 minus Jan. 1965 Diff.	r	t-Ratio	Diff. in Jan. Means, Control minus Exp. Diff.	r	t-Ratio	Diff. in May Means, Control minus Exp. Diff.	r	t-Ratio	df
Highest 1/4	Exper.	81	121.83	6.79	595.96	59.40	606.59	57.58	10.63	0.33	1.401	- 9.40	0.20	1.008	-15.54	0.13	1.714	80
"	Cont.	81	121.59	6.75	586.57	71.40	591.05	64.82	4.48	0.60	0.651							
2nd High 1/4	Exper.	116	105.93	3.51	532.23	64.33	525.46	63.59	-6.78	0.46	1.089	- 7.91	0.09	0.893	18.83	0.00	2.260*	115
	Cont.	116	106.01	3.62	524.32	76.35	544.28	63.05	19.97	0.38	2.733*							
3rd High 1/4	Exper.	90	94.11	3.66	491.43	52.83	491.16	59.40	-0.28	0.18	0.036	-14.61	0.02	1.781	11.21	0.04	1.246	89
	Cont.	90	94.10	3.77	476.82	57.72	502.37	58.31	25.54	0.34	3.627*							
Lowest 1/4	Exper.	80	80.35	6.74	432.43	58.47	444.25	58.75	11.83	0.34	1.564	23.96	0.24	2.981*	6.99	0.12	0.786	79
"	Cont.	80	80.84	6.87	456.39	57.54	451.24	60.60	-5.15	0.53	0.797							
Total Group	Exper.	365	100.88	15.34	514.33	81.89	516.98	82.01	2.65	0.65	0.738	- 2.79	0.47	0.628	7.30	0.44	1.640	364
"	Cont.	365	100.96	15.16	511.54	82.43	524.28	78.94	12.74	0.64	3.537*							

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

*Significant at 0.05 level (two-tailed test).

did the experimental students have a mean gain which was significant. Progress within the highest quarter and lowest quarter was moderate; however, in the middle quarters there was a slight loss. Somewhat the reverse pattern evolved for the control subgroup. There were strong gains in the two middle quarters (19.97 and 25.54), both significant, a slight gain in the top quarter, and a slight loss in the lowest quarter, neither significant.

The principal comparisons are between experimental subgroups in May. The mean difference of 7.30, favoring the controls, was not significant. Only at the second ability level was the control subgroup mean significantly higher than the experimental subgroup mean: 18.83, t-ratio of 2.260. At the highest quarter, the experimental subgroup mean surpassed the control subgroup mean (606.59 compared to 591.05; t-ratio of 1.714).

Theme Rating, Second Semester

Overall performance. Table XXVI shows the performance on theme in January and in May 1965 by the 365 matched pairs of students in the First Year Group. Analysis of theme data is limited to a comparison of the experimental subgroup and the control subgroup on each testing occasion. The basic fact in Table XXVI is that neither in January nor in May was there a significant mean difference between the two subgroups. The control subgroup was somewhat ahead--the mean differences were 0.26 (January) and 0.13 (May). Within each subgroup, the mean for females was higher than the mean for males. For experimentals, the mean differences were 1.33 in January and 0.52 in May. For controls, the corresponding figures were 1.01 and 0.58. It appears that during the second semester the males, although still significantly lower as a group, had narrowed the gap considerably.

Performance by ability quarters. One of the main functions of the analysis by ability levels is to see whether the overall analyses do, in fact, mask important characteristics present at one or more ability levels. The facts for the four ability levels presented in Table XXVII indicate that none of these levels is associated in any special way with the overall mean superiority of 0.13 held by the control over the experimental subgroup. Only in the highest quarter was the experimental mean higher than the control mean (by 0.40). None of the differences by ability level were significant.

TABLE XXVI

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

Subgroup	N	Degrees of Freedom	January Theme Mean	January Total S. D.	Difference in January Means Control minus Experimental Diff.	r	t-Ratio	May Theme Mean	May Total S. D.	Difference in May Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	365	364	9.95	2.38	0.26	0.26	1.730	9.63	2.27	0.13	0.23	0.864
Controls	365		10.21	2.35				9.76	2.21			
Exp.Males	134	133	9.11	2.55	0.46	0.27	1.790	9.30	2.26	0.09	0.26	0.372
Cont.Males	134		9.57	2.36				9.39	2.29			
Exp.Females	231	230	10.44	2.13	0.14	0.19	0.774	9.82	2.25	0.15	0.19	0.801
Cont.Females	231		10.58	2.26				9.97	2.13			

	January Diff. in Means Female minus Male Diff.	t-Ratio	May Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	1.33	5.329*	0.52	2.132*	363
Controls	1.01	4.026*	0.58	2.440*	363

*Significant at 0.05 level (two-tailed test).

TABLE XXVII

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON TOTAL THEME IN JANUARY 1965 AND MAY 1965 AT EACH OF FOUR Z-SCORE[1] LEVELS; MEANS, STANDARD DEVIATIONS, DIFFERENCES IN MEANS, AND t-RATIOS: COMBINED INSTITUTIONS

			Sept. 1964 Z-Score		Jan. 1965 Theme		May 1965 Theme		Difference in May Means Control minus Experimental			
Quarter	Subgroup	N	Mean	S. D.	Mean	S. D.	Mean	S. D.	Difference	r	t-Ratio	d.f.
Highest 1/4	Experimental	81	121.83	6.79	11.49	2.03	10.95	2.27	-0.40	0.35	1.321	80
"	Control	81	121.59	6.75	11.38	2.23	10.56	2.40				
Second High 1/4	Experimental	116	105.93	3.51	10.16	2.18	9.65	2.13	0.41	0.01	1.489	115
"	Control	116	106.01	3.62	10.63	2.24	10.05	2.03				
Third High 1/4	Experimental	90	94.11	3.66	9.63	2.15	9.38	1.91	0.14	0.04	0.498	89
"	Control	90	94.10	3.77	9.57	2.20	9.52	2.04				
Lowest 1/4	Experimental	80	80.35	6.74	8.45	2.19	8.49	2.17	0.25	0.20	0.842	79
"	Control	80	80.84	6.87	9.14	2.08	8.74	1.99				
Total Group	Experimental	365	100.88	15.34	9.95	2.38	9.63	2.27	0.13	0.23	0.864	364
"	Control	365	100.96	15.16	10.21	2.35	9.76	2.21				

---

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

COOP, First Year

Overall performance. Tables XXVIII and XXIX present data for the full freshman year: September 1964 through May 1965. These tables correspond to the series of tables for the first semester (Tables XVI to XXI) and the series for the second semester (Tables XXII to XXVII). Such a first semester, second semester, and combined semester report is based on the performance of the 365 matched pairs of students in the First Year Group.

In this section, attention will be focused upon gains for the nine-month period. On COOP, the gains in mean converted score, reported in Table XXVIII, September to May, were 3.90 (experimental) and 4.50 (control). For the COOP, the mean gain by the experimentals was accounted for about equally by each of the two semesters (1.91, first semester; 1.99, second semester).* For the controls, about two-thirds of the gain for the year occurred during the first semester (2.97 for the first semester, 1.53 for the second semester). Gains for each subgroup were significant each semester.

The September-to-May gains for the two sexes were similar: the smallest mean gain, 3.77, was by the experimental females; the largest mean gain, 4.63, was by the control females. Both of these gains were significant.

A recurring, striking aspect of the findings, the consistent superiority of female means over male means, is seen in the approximately three-point mean superiority in both September and May. The reader may note from Table XXVIII that this differential is approximately three-fourths as large as the mean freshman year gain on COOP (3.90 and 4.50). The question, How far are the males behind the females on COOP? might be answered by determining at what deferred testing date the COOP mean for a representative group of male students would be likely to equal the September COOP mean for a representative group of female students. The answer: apparently about seven months after September, or about March of the freshman year. To put it another way, the males lag behind the females by an interval of about seven months.

Performance by ability quarters by sex. Table XXIX, like Tables VIII, XI, and XIV (pages 61, 67, and 73) contains the most complete analysis made of the data by the investigators. It is a breakdown, by ability quarters, subgroup, and sex. The column headed "COOP, May minus September Mean" shows that September to May

---

*See Table XVI, page 77, and Table XXII, page 87, for these data.

TABLE XXVIII

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

Subgroup	N	Cooperative Eng. Test Converted Score September 1964		May 1965		Difference in Means, May minus September	r	t-Ratio	Difference in Sept. Means, Control minus Experimental			Difference in May Means, Control minus Experimental			Degrees of Freedom
		Mean	S. D.	Mean	S. D.				Diff.	r	t	Diff.	r	t	
Experimental	365	162.60	7.30	166.50	6.82	3.90	0.74	14.419*	0.19	0.80	0.791	0.79	0.55	2.283*	364
Control	365	162.79	7.40	167.29	7.09	4.50	0.71	15.601*							
Exp.Males	134	160.51	7.35	164.63	7.22	4.13	0.71	8.610*	0.13	0.77	0.291	0.28	0.44	0.431	133
Cont.Males	134	160.63	7.60	164.91	6.68	4.28	0.61	7.775*							
Exp.Females	231	163.81	6.99	167.58	6.33	3.77	0.73	11.593*	0.23	0.80	0.795	1.10	0.58	2.704*	230
Cont.Females	231	164.04	6.97	168.68	6.96	4.63	0.75	14.192*							

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	3.30	4.262*	2.95	4.055*	363
Controls	3.41	4.342*	3.77	5.042*	363

*Significant at 0.05 level (two-tailed test).

TABLE XXIX

PERFORMANCE ON COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION BY MATCHED PAIRS BY SEX IN SEPTEMBER 1964 AND MAY 1965 AT EACH OF FOUR Z-SCORE[1] LEVELS: COMBINED INSTITUTIONS

Subgroup	Level	N	Sept. Z-Score Rating Mean	S.D.	Sept. COOP Conv. Score Mean	S.D.	May COOP Conv. Score Mean	S.D.	COOP, May minus September Mean	r	t-Ratio	df	COOP, Female minus Male September Diff.	t-Ratio	May Diff.	t-Ratio	df
Exp.Male	Highest	17	119.24	4.44	171.06	4.44	174.06	4.93	3.00	0.20	2.020	16	0.59	0.511	-0.14	0.099	77
Exp.Female	1/4	62	122.71	7.15	171.65	4.05	173.92	5.15	2.27	0.70	4.802*	61					
Cont.Male	"	17	118.88	3.86	170.47	3.05	171.65	4.60	1.18	0.69	1.413	16	1.37	1.228	3.77	2.889*	77
Cont.Female		62	122.54	7.14	171.84	4.24	175.42	4.74	3.58	0.65	7.418*	61					
Exp.Male	2nd	42	105.86	3.78	164.57	3.12	167.40	5.14	2.83	0.53	4.137*	41	-0.16	0.255	0.24	0.249	114
Exp.Female	High 1/4	74	105.97	3.35	164.41	3.47	167.64	4.52	3.23	0.49	6.674*	73					
Cont.Male	"	42	105.83	3.85	164.76	3.95	167.24	5.27	2.48	0.27	2.788*	41	0.40	0.535	2.42	2.444*	114
Cont.Female		74	106.11	3.48	165.16	3.78	169.66	4.98	4.50	0.37	7.693*	73					
Exp.Male	3rd	31	94.52	3.89	159.90	4.14	162.52	5.76	2.61	0.45	2.663*	30	0.78	0.889	2.07	2.088*	88
Exp.Female	High 1/4	59	93.90	3.51	160.68	3.75	164.59	3.54	3.92	0.07	6.005*	58					
Cont.Male	"	31	94.29	4.13	160.71	4.73	165.03	5.20	4.32	0.25	3.889*	30	-0.64	0.718	0.39	0.368	88
Cont.Female		59	94.00	3.57	160.07	3.53	165.42	4.48	5.36	0.17	7.824*	58					
Exp.Male	Lowest	44	78.59	7.74	152.98	4.43	159.84	5.74	6.86	0.45	8.283*	43	1.27	1.203	1.60	1.909	78
Exp.Female	1/4	36	82.50	4.39	154.25	4.90	161.44	5.24	7.19	0.21	6.694*	35					
Cont.Male	"	44	79.15	7.98	152.84	5.22	160.00	5.94	7.16	0.26	6.912*	43	1.99	1.276	0.36	0.294	78
Cont.Female		36	82.92	4.39	154.83	3.66	160.36	4.63	5.53	0.05	5.671*	35					

[1] Combination of Cooperative English Tests: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

*Significant at 0.05 level (two-tailed test).

gains were largest in the lowest one-quarter, and that these relatively large gains were noted for both males and females, and for both experimentals and controls. That is, at this point, magnitude of gains was more a function of ability level than of treatment or sex. In the lowest quarter, the differences between COOP means in May and in September were of the order of six converted score points, compared to overall mean differences of about four converted score points.

CEEB, First Year

Overall performance. Table XXX shows that mean changes on CEEB during the freshman year were 32.79 and 41.47, broken down as follows:

Subgroup	First Semester*	Second Semester**	Full Year
Experimental	30.14	2.65	32.79
Control	28.72	12.74	41.47

Thus the increment for the year was attributable primarily to the first semester. Therefore, the progress reflected in the analysis for the full year was not much greater than the progress achieved by the end of the first semester. It appears that a fair amount of change occurred during the first semester whether instruction was present or not, while even the modest change during the second semester required the presence of instruction. No explanation for this reduced rate of gain during the second semester suggests itself to the investigators. In the discussion of Table XXVIII, page 98, the investigators pointed out that the COOP data showed a limited tendency for the change during the first semester to be greater than the change during the second semester.

Performance by ability quarters by sex. Table XXX indicates that for ability levels and sexes combined, the mean gain by the experimentals was 32.79 and by the controls 41.47. In Table XXXI, the same September-to-May CEEB data are presented by ability levels by sex. It is realized that as smaller groups are used for analysis, the resulting evidence is less reliable. The intention is to identify any suggestions of underlying facts which may be obscured by the larger

---

*See Table XVIII, page 81.
**See Table XXIV, page 91.

TABLE XXX

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

		College Entrance Examination Board Standard Rating				Difference in Means, May minus September			Difference in September Means, Control minus Experimental			Difference in May Means, Control minus Experimental			
		Sept. 1964		May 1965											
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t	Diff.	r	t	Degrees of Freedom
Experimental	365	484.20	83.25	516.98	82.01	32.79	0.64	8.931*	-1.38	0.81	0.513	7.30	0.44	1.640	364
Control	365	482.81	82.85	524.28	78.94	41.47	0.64	11.481*							
Exp. Males	134	467.49	85.47	501.47	77.43	33.98	0.70	6.204*	-1.13	0.78	0.230	-1.19	0.47	0.172	133
Cont. Males	134	466.36	86.65	500.28	77.51	33.92	0.58	5.169*							
Exp. Females	231	493.87	80.36	525.98	83.24	32.11	0.59	6.602*	-1.52	0.82	0.483	12.22	0.41	2.127*	230
Cont.Females	231	492.36	79.01	538.20	76.37	45.84	0.66	10.819*							

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	26.38	2.945*	24.51	2.774*	363
Controls	26.00	2.916*	37.92	4.536*	363

*Significant at 0.05 level (two-tailed test).

TABLE XXXI

PERFORMANCE ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TEST BY MATCHED PAIRS BY SEX IN SEPTEMBER 1964 AND MAY 1965 AT EACH OF FOUR Z-SCORE[1] LEVELS: COMBINED INSTITUTIONS

Subgroup	Level	N	Sept. Z-Score Rating Mean	S.D.	Sept. CEEB Stan. Rating Mean	S.D.	May CEEB Stan. Rating Mean	S.D.	CEEB, May minus Sept. Mean	r	t-Ratio	df	CEEB, Female minus Male September Diff.	t-Ratio	May Diff.	t-Ratio	df
Exp.Male	Highest	17	119.24	4.44	594.41	48.94	620.24	49.41	25.82	0.42	1.955	16	-11.30	0.886	-16.18	1.034	77
Exp.Female	1/4	62	122.71	7.15	583.11	45.14	604.06	57.75	21.05	0.24	2.557*	61					
Cont.Male	"	17	118.88	3.86	597.12	43.89	561.35	67.20	-35.76	0.04	1.815	16	-16.97	1.363	41.40	2.455*	77
Cont.Female		62	122.54	7.17	580.15	45.17	602.75	58.92	22.60	0.47	3.214*	61					
Exp.Male	2nd	42	105.86	3.78	515.21	40.94	519.00	55.03	3.79	0.35	0.433	41	- 5.13	0.641	10.12	0.819	114
Exp.Female	High 1/4	74	105.97	3.35	510.08	41.19	529.12	67.70	19.04	0.04	2.095*	73					
Cont.Male	"	42	105.83	3.85	512.36	37.46	535.50	65.58	23.14	0.22	2.177*	41	-10.02	1.172	13.77	1.127	114
Cont.Female		74	106.11	3.48	502.34	47.10	549.27	61.01	46.93	0.23	5.900*	73					
Exp.Male	3rd	31	94.52	3.89	449.26	43.96	504.26	45.89	55.00	0.03	4.825*	30	-11.58	1.120	-19.99	1.519	88
Exp.Female	High 1/4	59	93.90	3.51	437.68	47.16	484.27	64.33	46.59	0.17	4.866*	58					
Cont.Male	"	31	94.29	4.13	437.97	62.03	489.16	65.21	51.19	0.28	3.666*	30	7.71	0.671	20.15	1.561	88
Cont.Female		59	94.00	3.57	445.68	44.50	509.31	53.03	63.63	0.25	8.054*	58					
Exp.Male	Lowest	44	78.59	7.74	385.75	55.22	436.89	54.98	51.14	0.37	5.413*	43	13.22	1.110	16.36	1.236	78
Exp.Female	1/4	36	82.50	4.39	398.97	48.58	453.25	61.87	54.28	0.15	4.419*	35					
Cont.Male	"	44	79.15	7.98	391.93	57.77	450.89	65.58	58.95	0.52	6.383*	43	5.24	0.419	0.78	0.057	78
Cont.Female		36	82.92	4.39	397.17	51.25	451.67	53.90	54.50	0.21	4.888*	35					

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

*Significant at 0.05 level (two-tailed test).

picture. The column headed "CEEB, May minus September Mean" shows that the largest gains occurred in the two lowest ability quarters, that these gains were of similar magnitude for males and females, experimentals and controls, and that they were all significant. These mean advances varied from 46.59 (experimental females, third quarter) to 63.63 (control females, third quarter).

Even though there were only 17 control males in the top one-quarter and the results are therefore subject to sampling error, it is nevertheless noted that for these 17 males, there was a mean decrease in CEEB performance of 35.76 (t=1.815, not significant) from September to May. When the investigators analyzed the performance of this group by semesters, they discovered that the mean change during the first semester was -40.00, and during the second semester +4.23.* These 17 males happened to represent only three participating institutions. Two of the institutions had no control males in the top one-quarter who performed on all three testing occasions.

Theme Rating, First Year

Overall performance. According to Table XXXII, at the end of one full year of college the mean theme rating for the 365 control students was 9.76, and for the 365 experimental students, 9.63. The difference of 0.13 was not significant; a t-test for related measures yields a ratio of 0.864. When the analysis was made for the sexes separately, the results resembled those for the combined sexes (see lines 3, 4, 5, and 6 of Table XXXII). In the lower portion of Table XXXII is the comparison of theme ratings by sex. The findings for the experimental subgroup are similar to those for the control subgroup. At the outset (September 1964) the theme mean for the females was 0.63 higher than that for the males and at the end of the freshman year, the superiority was 0.52 for the experimental females and 0.58 for the control females. All these differences are significant. The fact that on the theme, the superiority of females over males was approximately the same at the end of the freshman year as it had been at the beginning of the year may be examined on the basis of the facts for each of the two semesters. A semester-by-semester analysis shows that at the beginning of the first semester the female mean theme score was 0.63 higher than that of the males both for the experimentals and for the controls. At the end of the first semester, a comparison of male and female means indicates the experimental female mean

---

*The analysis for the first and second semesters for these 17 males is derived from the computer output; it is not presented in any of the tables included in this report.

TABLE XXXII

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

			September Theme	September Total	Difference in September Means Control minus Experimental			May Theme	May Total	Difference in May Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Mean	S. D.	Diff.	r	t-Ratio	Mean	S. D.	Diff.	r	t-Ratio
Experimentals	365	364	9.23	2.01	0.00	1.00	0.000	9.63	2.27	0.13	0.23	0.864
Controls	365		9.23	2.01				9.76	2.21			
Exp.Males	134	133	8.83	2.14	0.00	1.00	0.000	9.30	2.26	0.09	0.26	0.372
Cont.Males	134		8.83	2.14				9.39	2.29			
Exp.Females	231	230	9.46	1.90	0.00	1.00	0.000	9.82	2.25	0.15	0.19	0.801
Cont.Females	231		9.46	1.90				9.97	2.13			

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.63	2.909*	0.52	2.132*	363
Controls	0.63	2.909*	0.58	2.440*	363

*Significant at 0.05 level (two-tailed test).

was 1.33 higher than that of the experimental males, while the control female mean was 1.01 above that of the control males.* Thus, midway in the freshman year, the females exhibited a greater superiority over the males than they did at either the beginning or the end of the freshman year. It therefore follows from all this that on the testing at the end of the second semester the males performed better relative to females than at the end of the first semester.

It was pointed out above that the obtained mean difference of 0.13 between experimental subgroup and control subgroup (Table XXXII) was not significant. How large a difference in means on total theme rating in May would be significant? For the given conditions of variability and correlation, it is possible to develop an estimate. The standard error of the mean difference is 0.146. Thus an obtained difference in means of 0.29 would be required for significance. A mean difference of 0.29 would have been present had 58 of the 365 control students received a rating one point higher than their actual rating.

The broadness of the scoring units for total theme rating may be noted in this connection. The distribution (N=1,978) of total theme ratings for both matched and unmatched students who wrote in May 1965 had a median of about 10. That is, a total theme rating of 10 had a percentile rank of 50. Near the center of this distribution, a shift of 1 point in theme rating is associated with a shift of 15 points in percentile rank:

Total Theme Rating	Percentile Rank
11	65
10	50
9	35

Total Theme Rating

September 1964			May 1965		
N	Mean	S. D.	N	Mean	S. D.
4,147	9.03	2.59	1,978	9.52	2.41

---

*See Table XX, page 84.

Performance by ability quarters by sex. Table XXXIII shows the theme evidence for September 1964 and May 1965, by subgroup, by sex, by quarter. In general, there was a direct relationship between quarters of ability (as derived from scores on two objective tests in September) and mean theme ratings in May. The right-hand portion of the table indicates the difference in means between females and males by subgroup by ability level in May. At the top of the ability distribution, the male means were higher than the female means--on the order of 0.40. At the center and lower part of the ability distribution, the female means were higher than the male means--from 0.42 to 0.83. None of the differences in means for females and males were significant; this is in part related to the fact that the N's were relatively small.

Two Year Sample

COOP Performance, First Semester

Overall performance. Table XXXIV contains COOP data based on the 122 matched pairs of students who completed the entire testing program. Performance is here reported over their first semester, September 1964-January 1965. The control subgroup made a greater gain than did the experimental subgroup (controls=2.59; experimentals=1.88), though the difference in performance between the two subgroups at the end of the semester (1.12 in favor of the controls) was not quite large enough to attain significance (attained t=1.956, 1.98 required for significance when N=122, degrees of freedom=121).

Performance by sex. In the middle portion of Table XXXIV the data are compared by sex: male experimentals compared with male controls and female experimentals with female controls. The males in the two subgroups attained about the same mean gains and end-of-semester means. The female controls, however, ended the semester significantly superior to the female experimentals (difference of 1.49, t=2.036).

The lower portion of Table XXXIV compares males with females within each subgroup. Here the females show a significant superiority over the males in both the experimental and the control subgroups at the beginning of the semester, that is, in September 1964 (3.65 for experimentals, 2.79 for controls). At the end of the first semester, the difference between males and females in the experimental subgroup (2.24) is no longer significant, owing to the greater gain of the males during the semester (male experimentals 2.77, female 1.37--see middle portion of the table). The female controls maintained their significant superiority over the male controls at the end of the first semester (3.26), having gained slightly more than the males (2.76 to 2.30--middle of the table).

TABLE XXXIII

PERFORMANCE ON THEME TOTAL BY MATCHED PAIRS BY SEX IN SEPTEMBER 1964 AND MAY 1965 AT EACH OF FOUR Z-SCORE[1] LEVELS: COMBINED INSTITUTIONS

Subgroup	Level	N	Sept. Z-Score Rating Mean	Sept. Z-Score Rating S.D.	September Theme Total Mean	September Theme Total S. D.	May Theme Total Mean	May Theme Total S. D.	Total Theme Rating Female minus Male September Diff.	Total Theme Rating Female minus Male September t-Ratio	Total Theme Rating Female minus Male May Diff.	Total Theme Rating Female minus Male May t-Ratio	df
Exp.Male	Highest	17	119.24	4.44	10.53	1.42	11.35	1.94	-0.09	0.203	-0.38	0.632	77
Exp.Female	1/4	62	122.71	7.15	10.44	1.73	10.97	2.26					
Cont.Male	"	17	118.88	3.86	10.53	1.42	10.94	2.21	-0.09	0.203	-0.41	0.616	77
Cont.Female		62	122.55	7.14	10.44	1.73	10.53	2.44					
Exp.Male	2nd	42	105.86	3.78	9.31	2.01	9.93	2.06	0.26	0.702	-0.44	1.070	114
Exp.Female	High 1/4	74	105.97	3.35	9.57	1.82	9.49	2.15					
Cont.Male	"	42	105.83	3.85	9.31	2.01	9.76	2.17	0.26	0.702	0.46	1.158	114
Cont.Female		74	106.11	3.48	9.57	1.82	10.22	1.92					
Exp.Male	3rd	31	94.51	3.89	8.58	1.74	9.00	1.87	0.20	0.494	0.58	1.361	88
Exp.Female	High 1/4	59	93.90	3.51	8.78	1.82	9.58	1.90					
Cont.Male	"	31	94.29	4.13	8.58	1.74	9.23	2.17	0.20	0.494	0.45	0.994	88
Cont.Female		59	94.00	3.57	8.78	1.82	9.68	1.95					
Exp.Male	Lowest	44	78.59	7.74	7.89	2.22	8.11	2.04	0.78	1.729	0.83	1.729	78
Exp.Female	1/4	36	82.50	4.39	8.67	1.65	8.94	2.24					
Cont.Male	"	44	79.14	7.98	7.89	2.22	8.55	2.12	0.78	1.715	0.42	0.950	78
Cont.Female		36	82.92	4.39	8.67	1.65	8.97	1.79					

---

[1] Combination of Cooperative English Test: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

TABLE XXXIV

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

		Cooperative Eng. Test Converted Score September		January		Difference in Means, Jan. minus September			Difference in Sept. Means Control minus Experimental			Difference in Jan. Means Control minus Experimental			
Subgroup	N	Mean	S. D.	Mean	S. D.		r	t-Ratio	Diff.	r	t	Diff.	r	t	Degrees of Freedom
Experimental	122	163.44	7.44	165.32	7.06	1.88	0.72	3.792*	0.41	0.79	0.939	1.12	0.60	1.956	121
Control	122	163.85	7.23	166.44	7.09	2.59	0.67	4.921*							
Exp.Males	44	161.11	7.20	163.89	6.94	2.77	0.66	3.112*	0.95	0.71	1.096	0.48	0.63	0.514	43
Cont.Males	44	162.07	7.82	164.36	7.20	2.30	0.52	2.034*							
Exp.Females	78	164.76	7.27	166.13	7.00	1.37	0.74	2.342*	0.10	0.82	0.216	1.49	0.57	2.036*	77
Cont.Females	78	164.86	6.66	167.62	6.76	2.76	0.76	5.220*							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	3.65	2.646*	2.24	1.690	120
Controls	2.79	2.067*	3.26	2.472*	120

*Significant at 0.05 level (two-tailed test).

COOP Performance, Second Semester

Overall performance. Table XXXV shows the second semester COOP results for the same 122 matched pairs whose first semester COOP performance was reported in the preceding table. The overall results are similar to those of the first semester. Both experimentals and controls made significant gains during the second semester (2.59 experimentals, t=5.041; 2.32 controls, t=4.939) as they had in the first semester. The difference in subgroup mean of 0.85 in favor of the controls at the end of the second semester was not significant (attained t=1.506, 1.98 required for significance with 121 degrees of freedom).

Performance by sex. Comparison between males and females in each of the two subgroups is presented in the middle and lower portions of Table XXXV. The gains made by both sexes in each subgroup on COOP during the second semester were significant (middle portion of table). Within the experimental subgroup, the differences in means in favor of the females (about two points ahead in both January and May) were not significant. Within the control subgroup, the difference in favor of the females (about three points ahead in both January and May) was significant.

COOP Performance, First Year

Overall performance. Table XXXVI, containing COOP results for the 122 matched pairs who completed the testing for the project, shows performance over the nine months from September 1964 to May 1965. The gains for both the experimental subgroup and the control subgroup (4.47 and 4.91) are significant. The mean difference between the performance of the two subgroups on the May testing, 0.85 in favor of the controls, is associated with a t-value of 1.506, with 1.98 required for significance. In short, the 122 members of the experimental subgroup performed substantially the same on COOP as their control counterparts at the beginning and at the end of their freshman year, despite the absence of instruction in freshman composition for the experimental students.

Performance by sex. Essentially the same picture emerges in subgroup comparison by sex. The experimental males made a slightly greater gain than the control males over the nine months: 5.50 and 4.80. The between-subgroups difference in favor of the male controls declined accordingly, from 0.93 at the beginning of the fall semester to 0.25 at the end of the spring semester. The rounding of decimals causes the apparent discrepancy. In the female pairs, on the other hand, the controls increased their advantage, moving from a superiority of 0.10 in September to a difference of 1.19 in May. All of these May differences were significant.

TABLE XXXV

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

Subgroup	N	Cooperative Eng. Test Converted Score January Mean	S. D.	May 1965 Mean	S. D.	Difference in Means, May minus January	r	t-Ratio	Difference in Jan. Means Control minus Experimental Diff.	r	t	Difference in May Means Control minus Experimental Diff.	r	t	Degrees of Freedom
Experimental	122	165.32	7.06	167.91	6.79	2.59	0.67	5.041*	1.12	0.60	1.956	0.85	0.58	1.506	121
Control	122	166.44	7.09	168.76	6.79	2.32	0.72	4.939*							
Exp.Males	44	163.89	6.94	166.61	6.64	2.73	0.73	3.582*	0.48	0.63	0.514	0.25	0.61	0.286	43
Cont.Males	44	164.36	7.20	166.86	6.35	2.50	0.71	3.125*							
Exp.Females	78	166.13	7.00	168.64	6.76	2.51	0.62	3.681*	1.49	0.57	2.036*	1.19	0.55	1.619	77
Cont.Females	78	167.62	6.76	169.83	6.80	2.22	0.71	3.803*							

	January Diff. in Means Female minus Male Diff.	t-Ratio	May 1965 Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	2.24	1.690	2.03	1.587	120
Controls	3.26	2.472*	2.97	2.352*	120

*Significant at 0.05 level (two-tailed test).

TABLE XXXVI

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

		Cooperative Eng. Test Converted Score September		May 1965		Difference in Means			Difference in Sept. Means Control minus Experimental			Difference in May Means Control minus Experimental			
Subgroup	N	Mean	S. D.	Mean	S. D.	May minus September	r	t-Ratio	Diff.	r	t	Diff.	r	t	Degrees of Freedom
Experimental	122	163.44	7.45	167.91	6.79	4.47	0.72	9.224*	0.41	0.79	0.939	0.85	0.58	1.506	121
Control	122	163.85	7.23	168.76	6.79	4.91	0.72	10.319*							
Exp.Males	44	161.11	7.20	166.61	6.64	5.50	0.72	6.897*	0.95	0.71	1.096	0.25	0.61	0.286	43
Cont.Males	44	162.07	7.82	166.86	6.35	4.80	0.66	5.283*							
Exp.Females	78	164.76	7.27	168.64	6.76	3.88	0.72	6.439*	0.10	0.82	0.216	1.19	0.55	1.619	77
Cont.Females	78	164.86	6.66	169.83	6.80	4.97	0.75	9.132*							

	September Diff. in Means Female minus Male		May 1965 Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	3.65	2.646*	2.03	1.587	120
Controls	2.79	2.067*	2.97	2.352*	120

*Significant at 0.05 level (two-tailed test).

It is instructive to arrange the four mean gains, September to May, in order of size.

Experimental Males	5.50
Control Females	4.97
Control Males	4.80
Experimental Females	3.88

It is clear that the absence of instruction did not handicap the males on COOP, and that the experimental females exhibited the smallest gain. The noteworthy fact is that because of the greater gain of the males within the experimental subgroup, the mean of the females was significantly higher than that of the males in September (3.65), but not in May (2.03). Within the control subgroup, the female mean exceeded the male mean significantly both in September and in May.

COOP Performance, Second Year

Overall performance. Table XXXVII presents the performance on COOP of the 122 persisting matched pairs who completed the May 1966 testing and compares their performance on that occasion with their performance in May 1965. The top two lines report the performance of the 122 experimentals and the 122 matched controls. Each subgroup performed essentially the same way in May 1966 as in May 1965 (mean declines of 0.67 for experimentals, 0.25 for controls). Those having had instruction did not perform significantly better on COOP than those not having instruction at the end of either the first year or the second year.

Performance by sex. Dividing the subgroups by sex and examining the results reveals that the males of both subgroups were similar in their May 1965 performance (mean differences of 0.25). On the May 1966 testing the control males showed a significant superiority (mean difference of 3.75). This difference results from a decline in performance by the experimental males (-2.07) coupled with a gain by the control males (1.43).

The bottom portion of Table XXXVII contains the array of COOP scores, May 1965 and May 1966, which shows most clearly the relationships between males and females within experimental subgroup and within control subgroup. In May 1965, the female mean was 2+ points more than the male mean in both subgroups, though this difference was significant only in the control subgroup. However, in May 1966 the experimental female superiority was 4+ points and the control female superiority almost disappeared.

TABLE XXXVII

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN MAY 1965 AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

Subgroup	N	Cooperative Eng. Test Converted Score May 1965 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May 1965 minus May 1966	r	t-Ratio	Difference in May 1965 Means Control minus Experimental Diff.	r	t	Difference in May 1966 Means Control minus Experimental Diff.	r	t	Degrees of Freedom
Experimental	122	167.91	6.79	167.24	8.93	-0.67	0.44	0.865	0.85	0.58	1.506	1.28	0.24	1.355	121
Control	122	168.76	6.79	168.52	7.89	-0.25	0.46	0.352							
Exp.Males	44	166.61	6.64	164.55	9.53	-2.07	0.68	1.948	0.25	0.61	0.286	3.75	0.42	2.645*	43
Cont.Males	44	166.86	6.36	168.30	7.46	1.43	0.51	1.361							
Exp.Females	78	168.64	6.76	168.76	8.19	0.12	0.25	0.110	1.19	0.55	1.619	-0.12	0.14	0.095	77
Cont.Females	78	169.83	6.80	168.64	8.12	-1.19	0.44	1.318							

	May 1965 Diff. in Means Female minus Male Diff.	t-Ratio	May 1966 Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	2.03	1.587	4.21	2.352*	120
Controls	2.47	2.548*	0.34	0.230	120

*Significant at 0.05 level (two-tailed test).

This situation reflects these main facts: female subgroups generally surpass male subgroups; between the two Mays the experimental males and the control females lost ground. In spite of the fact that control males were significantly superior to experimental males and that experimental females were significantly superior to experimental males in May 1966, the general finding for the sophomore year is an absence of mean gain for both the experimental and the control subgroups.

COOP Performance, Two Academic Years

Overall performance. Performance at the beginning of the freshman year (September 1964) and at the end of the sophomore year (May 1966) on COOP for the 122 matched pairs completing all project tests is the subject of Table XXXVIII. Over the span of two academic years, the experimental and control subgroups show similar significant gains: 3.80 for the experimentals (t=4.960), 4.66 for the controls (t=7.053). A t of 1.98 is sufficient for significance, degrees of freedom=121. The control subgroup started slightly higher (0.41) than the experimental subgroup and gained slightly more (0.86), but the end-of-two-year-difference (1.28) between subgroups is not significant (discrepancy is the result of rounding). As at the end of the first year and at the end of the second year, there was no significant difference in COOP scores between those who had composition instruction and those who did not when the two years are taken together.

Performance by sex. The control males out-gained the experimental males over the two academic years, 6.23 to 3.43. This difference in gain was great enough to provide the male controls a May 1966 mean significantly higher than the male experimental mean (difference of 3.75, attained t=2.645, 2.02 needed for significance with 43 degrees of freedom).

The female experimentals and controls had essentially the same mean gains--4.00 and 3.78. This similarity in mean performance by the female experimentals and controls existed both in September 1964 and May 1966.

The lower part of Table XXXVIII shows relationships of the sexes within each subgroup. The females were significantly superior to the males in September 1964: experimentals 3.65 and controls 2.79. In May 1966 the experimental females continued to show a significant superiority over the males (4.21) but the control females did not maintain a significant superiority over the control males (mean May 1966 difference=0.34). Within the control subgroup the two-year mean gain by the males was 6.23, and by the females 3.78.

TABLE XXXVIII

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

Subgroup	N	Cooperative Eng. Test Converted Score				Difference in Means May 1966 minus Sept.	r	t-Ratio	Difference in Sept. Means Control minus Experimental			Difference in May 1966 Means Control minus Experimental			Degrees of Freedom
		September Mean	September S. D.	May 1966 Mean	May 1966 S. D.				Diff.	r	t	Diff.	r	t	
Experimental	122	163.44	7.45	167.24	8.93	3.80	0.48	4.960*	0.41	0.79	0.939	1.28	0.24	1.355	121
Control	122	163.85	7.23	168.52	7.89	4.66	0.54	7.053*							
Exp.Males	44	161.11	7.20	164.55	9.53	3.43	0.66	3.130*	0.95	0.71	1.096	3.75	0.42	2.645*	43
Cont.Males	44	162.07	7.82	168.30	7.46	6.23	0.63	6.205*							
Exp.Females	78	164.76	7.27	168.76	8.19	4.00	0.32	3.887*	0.10	0.82	0.216	-0.12	0.14	0.095	77
Cont.Females	78	164.86	6.66	168.64	8.12	3.78	0.50	4.428*							

	September Diff. in Means Female minus Male		May 1966 Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	3.65	2.646*	4.21	2.548*	120
Controls	2.79	2.067*	0.34	0.230	120

*Significant at 0.05 level (two-tailed test).

Summary

Tables XXXIV through XXXVIII have presented COOP data for 122 matched pairs for four segments of the first two years of college and for the full two-year period. The general evidence concerning mean COOP gains is summarized in the following chart:

First Semester	
Exp. Cont.	GAIN

Second Semester	
Exp. Cont.	GAIN

First Academic Year	
Experimentals Controls	GAIN

Second Academic Year	
Experimentals Controls	NO GAIN

Two Academic Years	
Experimentals Controls	GAIN

There was a no-gain situation only during the sophomore year interval.

The main comparisons in this study deal with experimental and control subgroups. The evidence indicates that there was no significant difference between the scores on COOP of those who had had instruction in composition during the freshman year and those who had not had such instruction.

CEEB Performance, First Semester

Overall performance. Table XXXIX is the first in a series of five tables presenting the performance on CEEB of the 122 matched pairs who completed the entire testing program for the project. Table XXXIX concerns the performance at the beginning and at the end of the first semester in college, September 1964, and January 1965. The top two lines of Table XXXIX present the performance of the entire group of 122 pairs. Both the experimental subgroup and the control subgroup showed a significant gain in performance during the semester. The improvement of the control subgroup over the semester was not as great as that of the experimental subgroup (23.22 control, 34.86 experimental), with the result that in January the control mean was 15.20 less than the experimental mean, accounted for by the September disadvantage (3.56) and the smaller gain (11.64). This shows that of the group completing two years those not having composition instruction in the first semester scored significantly higher on CEEB at the end of that semester than did those having such instruction.

TABLE XXXIX

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in September Means, Control minus Experimental			Diff. in January Means, Control minus Experimental			
		September		January		Jan. minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	122	496.88	85.67	531.74	81.49	34.86	0.69	5.833*	-3.56	0.80	0.720	-15.20	0.46	1.869*	121
Control	122	493.32	86.07	516.54	90.76	23.22	0.61	3.269*							
Exp.Males	44	487.39	93.40	519.20	83.57	31.82	0.75	3.310*	-7.93	0.74	0.786	-31.43	0.34	2.233*	43
Cont.Males	44	479.45	88.82	487.77	76.71	8.32	0.51	0.661							
Exp.Females	78	502.23	80.50	538.81	79.42	36.58	0.65	4.779*	-1.09	0.84	0.207	-6.04	0.51	0.613	77
Cont.Females	78	501.14	83.47	532.77	93.99	31.63	0.66	3.738*							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	14.84	0.915	19.61	1.274	120
Controls	21.69	1.335	45.00	2.685*	120

*Significant at 0.05 level (two-tailed test).

Performance by sex. Significant gains were made during the first semester by the experimental males (31.82), the experimental females (36.58), and the control females (31.63). The gain of the control males (8.32) was not significant. The smaller gain by control males resulted in a significant superiority on January performance in favor of the experimental males (difference=31.43, t=2.233). Though the experimental females gained more than the control females (36.58 to 31.63), the difference on January means (6.04) was not significant.

Comparison of the performance of male and female members of the same subgroup shows a similar conclusion: that the females were somewhat, but not significantly, superior to the males in both subgroups in September (14.84 for experimentals, 21.69 for controls). When January comparisons are made, experimental and control females have increased their superiority over the corresponding males, the experimentals by 4.77 and the controls by 23.31. The superiority of control females over control males in January (45.00) was significant. That is, instruction in composition had a significantly greater effect on CEEB scores for women than for men at the end of the first semester.

CEEB Performance, Second Semester

Overall performance. Table XL is the second in the series presenting performance on CEEB of the 122 matched pairs completing the full project testing program. Data in Table XL are for the second semester, January 1965 and May 1965, test administrations. The experimental subgroup started the semester 15.20 points higher than the control subgroup, gained almost nothing during the semester, and ended the semester 4.31 points lower than the control subgroup. In contrast, the control subgroup showed a significant mean gain of 19.54. At the end of the second semester, there was no significant difference in the overall performance on CEEB between those who had had composition and those who had not.

Performance by sex. The mean gain of 19.54 by the control subgroup resulted from a mean gain of 22.93 by the 44 males and 17.63 by the 78 females, both of these being significant. The mean gain of 0.03 by the experimental subgroup resulted from a mean gain of 7.00 by the males and a mean loss of 3.90 by the females. The significant superiority of the experimental males over the control males on CEEB in January was lost in May, though the difference (15.50) was still in favor of the experimentals. The control females gained slightly more than the control males, although not enough to produce a significant difference between them.

TABLE XL

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in January Means, Control minus Experimental			Diff. in May Means, Control minus Experimental			
		January		May		May minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	January	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	122	531.74	81.49	531.77	74.62	0.03	0.66	0.006	-15.20	0.46	1.869	4.31	0.46	0.595	121
Control	122	516.54	90.76	536.08	78.77	19.54	0.69	3.159*							
Exp.Males	44	519.20	83.57	526.20	80.73	7.00	0.75	0.787	-31.43	0.34	2.233*	-15.50	0.58	1.443	43
Cont.Males	44	487.77	76.71	510.70	72.68	22.93	0.61	2.268*							
Exp.Females	78	538.81	79.42	534.91	70.74	-3.90	0.60	0.504	6.04	0.51	0.613	15.49	0.39	1.648	77
Cont.Females	78	532.77	93.99	550.40	78.47	17.63	0.69	2.244*							

	January Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	19.61	1.274	8.71	0.615	120
Controls	45.00	2.685*	39.70	2.732*	120

*Significant at 0.05 level (two-tailed test).

Comparison of the performance of males and females within subgroups consistently shows females superior to males, with significant superiority for control females both in January (45.00 points, t=2.685) and in May (39.70 points, t=2.732).

CEEB Performance, First Year

Third in the series reporting performance on CEEB for the 122 matched pairs completing the full project testing program is Table XLI, which gives the facts for the September 1964 and May 1965 administrations of CEEB, the beginning and the end of the students' freshman year in college.

Overall performance. Examination of the data indicates that the experimental subgroup began the year with a slight (3.56) advantage. Both subgroups made significant gains over the testing period, the experimentals improving by 34.89 points (t=5.507) and the controls by 42.76 points (t=6.934). Both gains were highly significant. The difference between the two subgroups at the end of the second semester was slight (4.31), indicating that instruction in composition had no significant effect on CEEB scores at the end of the year.

Performance by sex. Within the male group, the experimentals had a slightly higher mean gain than the controls: 38.82 to 31.25. Within the female group, the controls had a higher mean gain than the experimentals: 49.26 to 32.68. All four gains were significant. At the end of the first full year, the experimental males scored higher on CEEB than the control males (15.50), and the control females scored higher than the experimental females by an almost identical amount (15.49).

Comparison by sex within subgroups shows that while the females in both subgroups were superior to the males, the experimental females were not significantly so (8.71); the control females were significantly superior (39.70).

CEEB Performance, Second Year

Overall performance. Table XLII presents the performance on CEEB in May 1965 and May 1966, of the 122 matched pairs who completed the entire testing program. In examining this table it is important to remember that, during this second year of the study (1965-66), neither the experimental subgroup nor the control subgroup received instruction in freshman composition. During this twelve-month period each subgroup displayed about the same improvement, and in each case the improvement was significant. It is

TABLE XLI

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

		College Entrance Examination Board Standard Rating				Difference in Means May minus September			Diff. in September Means, Control minus Experimental			Diff. in May Means, Control minus Experimental			
		September		May											
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	122	496.88	85.67	531.77	74.62	34.89	0.63	5.507*	-3.56	0.80	0.720	4.31	0.46	0.595	121
Control	122	493.32	86.07	536.08	73.77	42.76	0.66	6.934*							
Exp.Males	44	487.39	93.40	526.20	80.73	38.82	0.76	4.144*	-7.93	0.74	0.786	-15.50	0.58	1.443	43
Cont.Males	44	479.45	88.82	510.70	72.68	31.25	0.57	2.696*							
Exp.Females	78	502.23	80.50	534.91	70.74	32.68	0.53	3.882*	-1.09	0.84	0.207	15.49	0.39	1.648	77
Cont.Females	78	501.14	83.47	550.40	78.47	49.26	0.71	6.993*							

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	14.84	0.915	8.71	0.615	120
Controls	21.69	1.335	39.70	2.732*	120

*Significant at 0.05 level (two-tailed test).

TABLE XLII

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN MAY 1965 AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

		College Entrance Examination Board Standard Rating				Difference in Means, May 1966 minus			Diff. in May 1965 Means, Control minus Experimental			Diff. in May 1966 Means, Control minus Experimental			
		May 1965		May 1966											
Subgroup	N	Mean	S. D.	Mean	S. D.	May 1965	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	122	531.77	74.62	554.62	81.65	22.85	0.66	3.878*	4.31	0.46	0.595	-2.52	0.54	0.367	121
Control	122	536.08	78.77	552.11	75.96	16.02	0.69	2.885*							
Exp.Males	44	526.20	80.73	548.52	77.53	22.32	0.81	2.962*	-15.50	0.58	1.443	-17.84	0.35	1.368	43
Cont.Males	44	510.70	72.68	530.68	72.39	19.98	0.68	2.262*							
Exp.Females	78	534.91	70.74	558.06	83.69	23.15	0.58	2.819*	15.49	0.39	1.648	6.13	0.64	0.796	77
Cont.Females	78	550.40	78.47	564.19	75.27	13.79	0.67	1.930							

	May 1965 Diff. in Means Female minus Male		May 1966 Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	8.71	0.615	9.54	0,616	120
Controls	39.70	2.732*	33.51	2.374*	120

*Significant at 0.05 level (two-tailed test).

noted that the mean CEEB gains during the sophomore year, on the order of 20 points, were about half as large as the mean gains during the freshman year (Table XLI). As in previous between-subgroup comparisons on CEEB, neither subgroup displayed a significant superiority on the May 1966 testing.

Performance by sex. The comparisons of males and females are presented, as in previous tables, in two ways. In the central portion of the table there is a between-subgroups comparison of males and females. Both experimental and control males improved significantly on CEEB during the second year; the males in the experimental subgroup performed somewhat better than the males in the control subgroup on both occasions. In contrast, the females of the control subgroup whose gain during the sophomore year was not significant, performed better both Mays than the females in the experimental subgroup, whose gain was significant. None of these mean differences between subgroups for the males or females on either testing occasion were significant.

In comparing males and females of a given subgroup, a noticeable dissimilarity between the subgroups is apparent. Within the experimental subgroup, the two sexes performed in essentially the same way in May 1965 and in May 1966, though the females displayed a slight advantage. Within the control subgroup, however, the females displayed a strong advantage on both occasions. The control females exhibited the highest mean scores of any of the four subgroups.

CEEB Performance, Two Academic Years

Overall performance. Table XLIII compares performance on CEEB of the 122 matched pairs who completed project testing through May 1966. The period between tests is four semesters, from September 1964 to May 1966--the beginning of the freshman year to the conclusion of the sophomore year. During this period both the experimental and control subgroups made significant improvement on CEEB. Both the experimental and control subgroups had beginning freshman means in the 490's, had two-year gains of about 58 points, and thus had ending sophomore means in the 550's. The gains were significant. Between-subgroup mean differences were not significant; composition instruction had no significant effect on CEEB scores at the end of the second year.

Performance by sex. When the evidence is analyzed by sex by subgroup, mean gains during the two years were roughly the same for males as for females. These four gains varied from 51.23 for control males to 63.05 for control females, all significant. Differences between control and experimental males, slightly in favor

TABLE XLIII

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

		College Entrance Examination Board Standard Rating				Diff. in Means, May 1966 minus			Diff. in September Means, Control minus Experimental			Diff. in May 1966 Means, Control minus Experimental			
		September		May 1966											
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	122	496.88	85.67	554.62	81.65	57.74	0.64	8.894*	-3.56	0.80	0.720	-2.52	0.54	0.367	121
Control	122	493.32	86.07	552.11	75.96	58.79	0.62	9.110*							
Exp.Males	44	487.39	93.40	548.52	77.53	61.13	0.70	5.876*	-7.93	0.74	0.786	-17.84	0.35	1.368	43
Cont.Males	44	479.45	88.82	530.68	72.39	51.23	0.55	4.299*							
Exp.Females	78	502.23	80.50	558.06	83.69	55.83	0.60	6.704*	-1.09	0.84	0.207	6.13	0.64	0.796	77
Cont.Females	78	501.14	83.47	564.19	75.27	63.05	0.66	8.354*							

	September Diff. in Means Female minus Male		May 1966 Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	14.84	0.915	9.54	0.616	120
Controls	21.69	1.335	33.51	2.374*	120

*Significant at 0.05 level (two-tailed test).

of the experimentals both in September 1964 and May 1966 (7.93 and 17.84), were not significant. Differences between the female experimentals and controls were slight.

When male-female performance within subgroups is examined, the females in each subgroup are found to be superior to the males, with the control females showing a significant superiority of 33.51 points (t=2.374) on their May 1966 means, the experimental females showing a non-significant superiority of 9.54.

Summary

Tables XXXIX through XLIII have presented data for 122 matched pairs of students on CEEB at each of four testing points in the first two years of college and for the full two-year period. The general data showing gains on CEEB are summarized in the following chart:

First Semester
Exp.
Cont. GAIN

Second Semester
Exp.
Cont. GAIN

First Academic Year
Experimentals
Controls GAIN

Second Academic Year
Experimentals
Controls GAIN

Two Academic Years
Experimentals
Controls GAIN

As the chart indicates, gain occurred between each set of testing occasions. Over no segment of time did the students in either subgroup fail to make some gain in performance on CEEB.

Theme Performance, First Semester

Theme performance for the 122 matched pairs available through the first two college years is presented in Tables XLIV through XLVIII. The data for the first semester appear in Table XLIV. Each table contains facts about theme ratings at the beginning and at the end of the given interval. Analysis of theme ratings is between subgroups for each testing occasion, and not within subgroups between two testing occasions. As has been pointed out previously, it is not meaningful to investigate change in theme performance over a specified interval (see page 42).

TABLE XLIV

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

			September Theme	September Total	Difference in September Means Control minus Experimental			January Theme	January Total	Difference in January Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Mean	S. D.	Diff.	r	t-Ratio	Mean	S. D.	Diff.	r	t-Ratio
Experimentals	122	121	9.31	1.87	0.00	1.00	0.000	10.20	2.24	0.22	0.29	0.912
Controls	122		9.31	1.87				10.42	2.08			
Exp.Males	44	43	9.07	1.64	0.00	1.00	0.000	9.57	2.42	0.36	0.23	0.822
Cont.Males	44		9.07	1.64				9.93	2.24			
Exp.Females	78	77	9.45	1.98	0.00	1.00	0.000	10.57	2.05	0.12	0.30	0.477
Cont.Females	78		9.45	1.98				10.69	1.94			

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.38	1.074	0.99	2.394*	120
Controls	0.38	1.074	0.76	1.950	120

*Significant at 0.05 level (two-tailed test).

Overall performance. In September 1964 the experimentals and the controls had identical means--9.31, because perfect matching on theme score rating was required. At the end of the first semester, the control mean was 0.22 higher than the experimental mean; this difference was not significant.

The correlation data are of interest. The matching procedures produced an r of 1.00 between the September theme ratings of the 122 experimentals and their 122 matched controls. The r at the end of the first semester was 0.29. This substantial decrease in the coefficient of correlation stems from the following: the relatively small number of rating values (the rating scale ranged from 2 to 18), the factors of unevenness in achievement within the matched pairs, and the unreliability present in the theme ratings. For the two objective tests, COOP and CEEB, the between-subgroup correlations were, in September, lower than 1.00, but in January higher than 0.29. (Additional discussions of correlation data will be found on page 50.)

Performance by sex. The analyses by sex show that in January the mean for control males was higher (0.36), though not significantly higher, than that of experimental males. Control females were even less superior (0.12) to experimental females. In September the mean for females was 0.38 higher than the mean for males in both subgroups. In January, the differences were 0.99 for the experimental subgroup and 0.76 for the control subgroup. Of these four female-male mean differences, only the 0.99 was significant.

Theme Performance, Second Semester

Overall performance. Table XLV contains the end-of-first-semester theme data which were reported in the preceding Table (XLIV) and also the end-of-second-semester theme data for the 122 pairs who finished the full project testing program. At the end of the freshman year, the mean theme ratings for the experimental subgroup and the control subgroup were only 0.06 apart (9.93 experimental, 9.87 control). That is, as measured in this project, theme ratings of students finishing two years of college who had completed freshman composition were not significantly different at the end of two semesters than theme ratings of students who had had no freshman composition.

Performance by sex. The experimental-control similarity was present for both the males (mean difference of 0.20 in favor of the experimentals) and the females (mean difference of 0.00). The disparity between the mean for females and the mean for males was less at the end of the second semester than at the beginning of the semester. Within the experimental subgroup, the mean difference,

TABLE XLV

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

			January		Difference in January Means Control minus Experimental			May		Difference in May Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Theme Mean	Total S. D.	Diff.	r	t-Ratio	Theme Mean	Total S. D.	Diff.	r	t-Ratio
Experimentals	122	121	10.20	2.24	0.22	0.29	0.912	9.93	2.25	-0.06	0.15	0.251
Controls	122		10.42	2.08				9.87	2.14			
Exp.Males	44	43	9.57	2.42	0.36	0.23	0.822	9.70	2.25	-0.20	0.08	0.444
Cont.Males	44		9.93	2.24				9.50	2.21			
Exp.Females	78	77	10.57	2.05	0.12	0.30	0.477	10.07	2.24	0.00	0.17	0.040
Cont.Females	78		10.69	1.94				10.07	2.07			

	January Diff. in Means Female minus Male		May Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.99	2.394*	0.37	0.842	120
Controls	0.76	1.950	0.57	1.431	120

*Significant at 0.05 level (two-tailed test).

females minus males, dropped from 0.99 (which was significant), to 0.37 (which was not significant); within the control subgroup, from 0.76 to 0.57 (neither significant). Thus there is a suggestion that at the end of the second semester the males were in a better position relative to the females than they had been at the beginning of that semester.

Theme Performance, First Year

Overall performance. The September and May data for the 1964-65 year in Table XLVI complete the analysis of theme performance of the first two semesters. The key fact is that the experimental subgroup and the control subgroup, which had identical means in September and a differential of only 0.22 in the middle of the year (Table XLV), also had near-identical means in May (difference of 0.06 in favor of experimentals).

Performance by sex. There was also close agreement between the September data and the May data as regards the superiority of females over males: 0.38 for both in September, 0.37 (experimentals) and 0.57 (controls) in May. None of these differences were significant (approximately 0.80 would be required for significance).

Theme Performance, Second Year

Overall performance. Table XLVII covers the sophomore year. It is unique in that it represents the period during which neither the experimental nor the control subgroup received formal instruction in freshman composition. In May 1966 the experimental mean (9.70) was 0.31 higher than the control mean (9.39), a difference too small for significance, though slightly larger than the mean difference at the end of the first year (0.06).

Performance by sex. The analysis by sex yielded results similar to those of the total group, the experimental males scoring 0.43 over control males, experimental females 0.24 over control females. Corresponding mean differences at the end of the first year had been 0.20 and 0.00. The most interesting disclosure of Table XLVII is that the experimental males, who were 0.37 behind the experimental females at the end of the freshman year, were 0.11 ahead of the experimental females at the end of the sophomore year, and that the superiority of the control females over the control males dropped from 0.57 at the end of the freshman year to 0.08 at the end of the sophomore year.

TABLE XLVI

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

			September		Difference in September Means Control minus Experimental			May		Difference in May Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Theme Mean	Total S. D.	Diff.	r	t-Ratio	Theme Mean	Total S. D.	Diff.	r	t-Ratio
Experimentals	122	121	9.31	1.87	0.00	1.00	0.000	9.93	2.25	-0.06	0.15	0.251
Controls	122		9.31	1.87				9.87	2.14			
Exp.Males	44	43	9.07	1.64	0.00	1.00	0.000	9.70	2.25	-0.20	0.08	0.444
Cont.Males	44		9.07	1.64				9.50	2.21			
Exp.Females	78	77	9.45	1.98	0.00	1.00	0.000	10.07	2.24	0.00	0.17	0.040
Cont.Females	78		9.45	1.98				10.07	2.07			

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.38	1.074	0.37	0.842	120
Controls	0.38	1.074	0.57	1.431	120

3

TABLE XLVII

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN MAY 1965 AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

			May 1965 Theme Total		Difference in May 1965 Means Control minus Experimental			May 1966 Theme Total		Difference in May 1966 Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Mean	S. D.	Diff.	r	t-Ratio	Mean	S. D.	Diff.	r	t-Ratio
Experimentals	122	121	9.93	2.25	-0.06	0.15	0.251	9.70	2.63	-0.31	0.16	1.012
Controls	122		9.87	2.14				9.39	2.59			
Exp.Males	44	43	9.70	2.25	-0.20	0.08	0.444	9.77	2.88	-0.43	0.24	0.790
Cont.Males	44		9.50	2.21				9.34	2.95			
Exp.Females	78	77	10.07	2.24	0.00	0.17	0.040	9.66	2.48	-0.24	0.09	0.654
Cont.Females	78		10.07	2.07				9.42	2.36			

	May 1965 Diff. in Means Female minus Male		May 1966 Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.37	0.842	-0.11	0.212	120
Controls	0.57	1.431	0.08	0.167	120

How, then, did the mean theme ratings at the end of the sophomore year compare with those at the end of the freshman year? First, those who had no freshman composition performed on both occasions about as well as those who did have freshman composition. Second, the slight superiority of females over males on theme rating, evident at the end of the freshman year (and, on nearly all objective tests at all testing sessions--see Tables XXXIV-XXXVIII for COOP, Tables XXXIX-XLIII for CEEB), did not exist at the end of the sophomore year.

Theme Performance, Two Academic Years

Table XLVIII depicts theme performance at the beginning and at the end of the first two college years. Analyses of theme data are always between groups as of a specified date and never within groups between two specified dates.

Overall performance. The data in this table are very similar to those of Table XLVII, the data for May 1966 being identical. By May 1966 the experimental subgroup scored a little higher (0.31), though not significantly higher than the control subgroup; both subgroups had the same mean in September 1964.

Performance by sex. In September 1964 males and females had identical mean theme ratings within subgroups (experimentals each 9.07, controls each 9.45). In May 1966 experimental males exceeded control males by a difference of 0.43; experimental females exceeded control females by 0.24. Neither of these differences is significant. Experimental females were 0.11 ahead of experimental males, control females 0.08 ahead of control males.

Conclusions for the data in Table XLVIII are the same as those for Table XLVII: after two years of college, students not receiving instruction in freshman composition performed as well on the theme as students who had received such instruction. Females, who were superior to males, not significantly on theme (0.38), significantly on objective tests, in September 1964, were negligibly ahead (experimentals 0.11, controls 0.08) in May 1966. In September 1964 the mean theme ratings for females were superior (not significantly) to those of males--0.38 for each subgroup. Two academic years later, the scores of the females and the males were even closer together.

TABLE XLVIII

THE PERFORMANCE OF 122 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: COMBINED INSTITUTIONS

			September Theme Total		Difference in September Means Control minus Experimental			May 1966 Theme Total		Difference in May 1966 Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Mean	S. D.	Diff.	r	t-Ratio	Mean	S. D.	Diff.	r	t-Ratio
Experimentals	122	121	9.31	1.87	0.00	1.00	0.000	9.70	2.63	-0.31	0.16	1.012
Controls	122		9.31	1.87				9.39	2.59			
Exp.Males	44	43	9.07	1.64	0.00	1.00	0.000	9.77	2.88	-0.43	0.24	0.790
Cont.Males	44		9.07	1.64				9.34	2.95			
Exp.Females	78	77	9.45	1.98	0.00	1.00	0.000	9.66	2.48	-0.24	0.09	0.654
Cont.Females	78		9.45	1.98				9.42	2.36			

	September Diff. in Means Female minus Male		May 1966 Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.38	1.074	-0.11	0.212	120
Controls	0.38	1.074	0.08	0.167	120

Summary

As the investigators do not believe the theme performances can be legitimately examined for gains, the summary below indicates the relative position of the subgroups of the 122 pairs who completed the full two academic years of the investigation on theme performance at each testing point. The following chart is based upon Tables XLIV through XLVIII.

Beginning of First Semester		End of First Semester	
Exp. Cont.	SAME	Exp. Cont.	CONTROL HIGHER

End of Second Semester		End of Fourth Semester	
Exp. Cont.	EXPERIMENTAL HIGHER	Exp. Cont.	EXPERIMENTAL HIGHER

The difference in favor of the controls at the end of the first semester was great enough to be significant. On the May 1965 and May 1966 testing occasions, the difference, in favor of the experimentals, was not significant.

SUMMARY, CONCLUSIONS, AND OBSERVATIONS

Overall Findings

The present study tested the hypothesis that the writing performance of students enrolled in a college composition sequence is not significantly different from the writing performance of comparable students not enrolled in a college freshman composition course when the two subgroups have attended college for an equal length of time. The total sample was composed of a representative sample from each of five state universities.

Students who received instruction excelled students who had not received instruction at the end of the first semester, and tended to surpass them at the end of the second semester; at the end of the fourth semester the two groups performed about the same. The data are summarized in the tabulation below. In this presentation C signifies that the control subgroup, the students who received instruction, had the higher obtained mean. E signifies that the experimental subgroup, the students who did not receive instruction, had the higher obtained mean. C* signifies that the controls had a significant superiority, E* that the experimentals had a significant superiority.

Test	Number of Matched Pairs	January 1965 (end of 1st Semester)	May 1965 (end of 2nd Semester)	May 1966 (end of 4th Semester
COOP	597	C*		
CEEB	597	C		
Theme	597	C*		
COOP	365	C*	C*	
CEEB	365	E	C	
Theme	365	C	C	
COOP	122	C	C	C
CEEB	122	E	C	E
Theme	122	C	E	E

It is evident that at the end of the first semester the 597 students who had received instruction in freshman composition performed significantly better than those who had not--on COOP and Theme performance. At the end of the second semester, the 365 control students performed significantly better than their experimental matches on COOP, but essentially the same on CEEB and theme. In May 1966 the 122 control students and their experimental matches performed in essentially the same way.

Experimental and control subgroups were compared 18 times. These facts concerning the performance of the subgroups may be summarized both from the point of view of the number of times a given subgroup excelled the other and from the point of view of the performance of the subgroups on each of the three test instruments. Summarizing first from the point of view of the number of times one subgroup excelled the other:

1. In the 18 comparisons, 4 showed a significant difference between the subgroups, in all 4 the control subgroup excelled.

2. In the 14 comparisons which did not reveal a significant difference between the two subgroups, the higher obtained mean was achieved 9 times by the controls and 5 times by the experimentals.

3. At no testing point was there a significant difference between the subgroups (122 pairs) who completed two years of college. The controls attained a superiority in observed mean 5 times, the experimental subgroup 4 times.

Summarizing next in terms of the performance of the subgroups on the testing instruments employed:

1. COOP - The control subgroup mean was significantly higher than the experimental subgroup mean in 3 comparisons and somewhat higher in 3 comparisons.

2. CEEB - In none of the 6 comparisons was there a significant difference between the means of the two subgroups: the obtained means favored each subgroup 3 times.

3. Theme - The control subgroup mean was significantly higher once. On the other 5 occasions, the obtained mean favored the controls 3 times and the experimentals twice.

In terms of testing instruments, then, COOP yielded superiority for the controls; the theme, if it favored either subgroup, favored the controls; the CEEB evidence suggested essential similarity between the control subgroup and the experimental subgroup. In general, COOP denied the hypothesis, the theme leaned slightly toward denial, and CEEB neither confirmed nor denied the hypothesis.

Do college students who have had formal course work in freshman English composition perform better on tests related to writing than comparable students who have not had the formal course work? Evidence of performance on the tests used in this study has shown that the answer at the end of the first semester is "Yes," at the end of the second semester, a qualified "Yes," and at the end of the fourth semester, "No." The two subgroups of students who finished the two years appear to be substantially equal.

The design of this study has combined the performance of students from several universities, each of which had a freshman program somewhat different from each of the others. At each university, members of the control subgroup received their instruction from several different instructors. Each instructor interpreted the official syllabus of his institution in his own way. None of the evaluative instruments employed in this study was attuned to a particular instructor or a particular university. Thus such differences as appear between control and experimental subgroups reflect the common elements which are present independently of the unique standards and qualities stressed in a particular program or class. Because the obtained differences are based on such common elements, they constitute only a partial basis for evaluating a program at one of the cooperating universities.

Findings by Sex

In both the Interim Report and the present study, the investigators have been interested in the relationship between sex of students and their performance on tests related to composition. The investigators' belief that there is such a relationship influenced the matching procedures, which employed sex as one of the criteria for matching students. Throughout the present report, except when numbers were so small that a division by sex would have led to confusion, performance has been reported both for total subgroups and for the male and female components of those subgroups. The present discussion summarizes the findings concerning performance by sex.

Superiority of female performance. It is useful at the outset to examine the distribution of performance on a national test, as it was in terms of such a distribution that the samples were selected. Data are at hand showing the performance of a normative sample of 882,080 high school seniors, and of 703 freshman males and 1,075 freshman females enrolled at the University of Northern Iowa in the fall of 1966. As described on page 39, the selection of students for the current investigation was in terms of distributions of scores on American College Testing Program, a separate distribution for each sex. The following tabulation is an illustration of the differences among the three distributions of ACT scores.

ACT English Standard Score	Percentile Rank National College-Bound High School Seniors	Percentile Rank, UNI Freshman Males N=703	Percentile Rank, UNI Freshman Females N=1,075
33			99
31	99		99
30	99	99	98
24	81	76	56
21	59	43	22
14	13	3	0
11	5	0	0

Among the illustrative standard scores included are, for the females, the highest score (33), the lowest score (14), and the score half-way between (24), not necessarily the median. For the males, the highest standard score was 30, the lowest 11, and the score half-way between, 21. The superiority of the female is clear when one notes that the middle female score (24) has a percentile rank of 56 in the UNI female score distribution, and of 76 in the UNI male distribution. The middle score of the male range, 21, has a corresponding percentile rank of 43 in the UNI male distribution and 22 in the UNI female distribution. In short, the typical male performs less well on the ACT English test than does the typical female.

The differences in performance of males and females illustrated in the tabulation above are reflected in the male-female performance of the 1,040 matched pairs (422 males, 618 females), from the cooperating universities in September 1964, the outset of the study. The performance by sex, and the t-ratios for differences in mean performances, are shown below.

Subgroup	N	Sex	Variable	Mean	S. D.	Female minus Male	t-Ratio	d.f.
Exp.	422	M	COOP	159.12	7.55	4.16	8.877*	1,038
Exp.	618	F	COOP	163.28	7.32			
Cont.	422	M	COOP	159.07	7.48	4.30	9.307*	1,038
Cont.	618	F	COOP	163.37	7.19			
Exp.	422	M	CEEB	445.83	85.41	44.11	8.368*	1,038
Exp.	618	F	CEEB	489.94	81.99			
Cont.	422	M	CEEB	446.67	85.01	42.84	8.170*	1,038
Cont.	618	F	CEEB	489.51	81.52			
Exp.	422	M	Theme	8.37	2.26	1.05	7.688*	1,038
Exp.	618	F	Theme	9.42	2.09			
Cont.	422	M	Theme	8.37	2.26	1.05	7.688*	1,038
Cont.	618	F	Theme	9.42	2.09			

*Significant at 0.01 level (two-tailed test).

The superiority of females reflected in the above tabulation may at first glance seem at variance with conclusions reached by other investigators. In an attempt to probe the male-female performance in writing, the investigators examined the studies of Hunt (1965), Riling (1965), O'Donnell and Griffin (1967), and Loban (1966). All of these are studies of the development of syntactic control in children. The authors comment, directly or indirectly, on the differences in performance between males and females. Hunt indirectly, and O'Donnell and Griffin directly, indicate that though males are initially at a disadvantage, the "gap" closes as the students grow older. After indicating that "In writing, . . . , girls in Grades 3 and 5 appeared to be superior to boys" (p. 96), O'Donnell and

Griffin state that "In the seventh grade . . . the relative positions of the sexes were clearly reversed on the scales taken to indicate syntactic skill." (p. 96). Hunt's tabulations of T-units shorter than nine words and longer than twenty words (pp. 28 and 31) would support somewhat the same conclusion, as at the eighth and twelfth grade levels the boys write fewer T-units shorter than nine words than do the girls, and at the twelfth grade level they write more T-units longer than twenty words than do the girls. The number of words per T-unit is used by Hunt as an index of maturity in writing.

Riling (p. 87) and Loban (p. 90), on the other hand, assert that the best boys do better than the best girls, but clearly imply that in general the boys do worse than the girls, and that the worst boys are worse than the worst girls. In an attempt to check on the latter point, the present investigators examined performance on COOP of 4,190 students for whom data were available in the present study (see discussion, Table VIII, page 60). The investigators stated ". . . that only at the top edge of the distribution (the top 2 percent) do males equal or excel females. At the bottom edge of the distribution the reverse is true; the male group falls below the females." (p. 62). This conclusion is supported by examination of the quarters-by-sex tables for the test instruments (Tables VII, p. 59; XI, p. 67; XIV, p. 73; XXIX, p. 99; XXXI, p. 102; XXXIII, p. 107. These tables show, both in the means for the males and in the means for the females, and in the proportion of each sex in the top and the bottom quarters, that the male distribution tends to the lower quarters, the female to the higher quarters. In short, the statements by Riling and Loban are consistent with the findings of this study--that while the very best writer, or performer on tests related to writing, may be a male, the mean of the males is lower than the mean of the females.

It is important, also, to remember that the O'Donnell and Griffin, Hunt, Riling, and Loban studies concerned the syntactic virtuosity of the children studied. That is, the investigators inquired into the kinds of syntactic structures the students employed. It is on such measures that they estimate the linguistic maturity of their subjects. In the present study, such matters were hardly noticed. Neither the objective tests nor the theme evaluations involved analyses of sentence structure, T-units, or other syntactically defined linguistic entities. It is not possible to compare the two kinds of studies in any direct way. One may theorize that the readers of the compositions in the present study reacted in some degree to the syntactical resourcefulness displayed in the papers. However, the degree to which such resourcefulness, or particular manifestations of such resourcefulness, influenced the readers' decisions is, of course, unknown. For the most part, the readers were probably not aware of the syntactic elements in the papers.

The relationship of male to female means presented in the various tables throughout this study is summarized in the following tabular representation, in which F stands for a female mean not significantly greater than a male mean, an M a male mean not significantly greater than a female mean, and an F* or M* a significant difference.

			Sept. 1964		Jan. 1965		May 1965		May 1966	
Test	N Pairs[a]	M/F	Exp.	Cont.	Exp.	Cont.	Exp.	Cont.	Exp.	Cont.
COOP	597	470/724	F*	F*	F*	F*				
CEEB	597	470/724	F	F	F	F				
Theme	597	470/724	F*	F*	F*	F*				
COOP	365	268/462	F*	F*	F*	F*	F*	F*		
CEEB	365	268/462	F*	F*	F*	F*	F*	F*		
Theme	365	268/462	F*	F*	F*	F*	F*	F*		
COOP	122	88/156	F*	F*	F	F*	F	F*	F*	F
CEEB	122	88/156	F	F	F	F*	F	F*	F	F*
Theme	122	88/156	F	F	F*	F	F	F	M	F

[a] The reader should remember that each succeeding N represents the persisting members of the preceding N.

In the 54 comparisons presented, males out-performed females on one occasion and the females were significantly superior on 35--the males were never significantly superior. Thus females performed significantly better than males on nearly two-thirds of the testing occasions. The control females were significantly superior on 19 occasions. There is therefore little doubt that in the population studied, the females were superior on tests related to writing ability. Twenty-six of the 30 comparisons made through May 1965 show the females significantly above the males. In the group which completed testing through January 1965 (N=597 pairs), the females were significantly superior on all theme and COOP comparisons, superior by an amount short of significance on CEEB only.

Gains by sex. The summary presented above indicates that, whatever may be the situations at the extremes of test score distributions, the mean performance of females is consistently superior to that of males in the group which persists through the first year of college. Do the males ever "catch up"? The evidence to answer that question is apparently not in, but there is a suggestion in the performance of the 88 males and 156 females who completed testing through the second year of college. Though their performance does not disclose male superiority, there is a slight indication that female superiority is decreasing, as the experimental males on the May 1966 theme attained a mean somewhat, but not significantly, greater than the female mean. Of the 24 comparisons presented for these 244 students, 9 show the females significantly superior, none shows the males significantly superior. Whether this trend would continue, and whether it indicates a leveling of performance or the persistence of the better male pairs would be difficult to say. Another possibility might be a difference in males and females in response to the test situation. In any event, one may firmly conclude by reiterating a statement made in the Interim Report ". . . that in investigations concerning competence in composition the ratio between sexes must be taken into account in the groups whose performance is being studied." (p. 65).

Thus far all comparisons of male and female performance reflect the relationship between the sexes on specific testing occasions. They may be summarized by saying that the mean of the females was superior at the beginning of the freshman year and remained so during the freshman year. A different question is whether one sex appears to benefit from instruction more than the other. The tabulation below presents a summary of the gains on the objective tests for the 365 pairs completing the first full year of the study. Asterisks indicate significant mean gains. The difference between male and female performance on themes is included only for completeness (the reason for this has been discussed on page 42). In the theme portion of the tabulation, asterisks represent significant differences between male and female means on theme performance as of the testing date.

Test	Subgroup	Sex	N	Sept.-Jan. Mean Gain	Jan.-May Mean Gain
COOP	Exp.	M	134	2.26*	1.87*
COOP	Exp.	F	231	1.71*	2.06*
COOP	Cont.	M	134	2.68*	1.60*
COOP	Cont.	F	231	3.14*	1.49*
CEEB	Exp.	M	134	26.97*	7.01
CEEB	Exp.	F	231	31.98*	0.13
CEEB	Cont.	M	134	20.52*	13.40*
CEEB	Cont.	F	231	33.48*	12.36*

				Jan. Theme	Diff. in Mean (F-M)	May Theme	Diff. in Mean (F-M)
Theme	Exp.	M	134	9.11	1.33* (F)	9.30	0.52* (F)
Theme	Exp.	F	231	10.44		9.82	
Theme	Cont.	M	134	9.57	1.01* (F)	9.39	0.58* (F)
Theme	Cont.	F	231	10.58		9.97	

On COOP both sexes in both subgroups made significant mean gains both semesters. During the first semester the experimental males gained slightly more than the experimental females; during the second semester the experimental females gained slightly more than the experimental males. Among the controls, the situation was reversed.

On CEEB significant mean gains were achieved by the control males and females both semesters, but by the experimental males and females only the first semester. As with COOP, there was inconclusive evidence concerning the possible superiority of one sex over the other in gains.

Theme performance. On theme performance, the females scored means significantly greater than those of the males both in January and in May of the freshman year. Following are summary statements concerning sex and performance on tests related to competence in written composition:

1. Though the best performance may be by a male, the mean of female performances is consistently higher, frequently significantly higher, than the mean for males.

2. This being true, in investigations concerning student performance on tests related to composition the percentage of males to females must be taken into account.

3. The disparity of performance between the sexes in tests assessing ability in composition persists at least through the sophomore year of college, with the possible exception of actual writing performance.

Performance by ability quarters, by sex. It is informative to examine performance by ability quarters, and within quarters, by sex. Performance on each of the three testing instruments is first discussed separately, followed by a summary statement.

1. COOP - For the 597 matched pairs of students who completed the first semester, analysis of the January 1965 COOP scores for the two subgroups showed that the control mean was significantly higher than the experimental mean. An analysis by quarters of ability showed that at all four ability levels the control mean surpassed the experimental mean, significantly so at the second highest level. When the means for males and females within treatments were compared at each ability level, it was found that in the highest and lowest levels the female mean exceeded the male mean. The COOP data by ability level did not reveal any instances of superiority of experimentals over controls or of males over females.

2. CEEB - On CEEB, the evidence for the 597 matched pairs showed similar January means for the experimental subgroup and the control subgroup in terms of the score scale involved. The noticeable exception to the overall evidence was in the lowest quarter, in which the control mean was significantly higher than the experimental mean. The analysis by sex by ability quarters showed that the control females surpassed the control males significantly in the top quarter, and the experimental females surpassed the experimental males significantly in the lowest quarter. In each of the two middle quarters, the females and males were about equal.

3. Theme - On the theme, a significant mean difference favoring the controls was found at the end of the first semester in the comparison for the complete subgroups and in the comparison for the lowest quarter. In the top quarter and the third quarter, the subgroup means were basically the same. The general finding concerning males

and females was that the females definitely excelled in the lowest two quarters, more so within the experimental subgroup than within the control subgroup.

Summary - If one looks at the analysis by sex and by ability levels of the evidence on the three testing instruments, the lowest quarter in ability seems to be unique. Here is to be found the most frequently recurring indication that control subgroup means are higher than experimental subgroup means and that female means are higher than male means.

This summarization has been based upon data through the first semester. It is for this sample of 597 matched pairs, the largest sample available for this purpose, that it is most defensible to utilize the finer analyses by ability and sex.

Number of Students

The amount of difference between treatments in treatments studies in college freshman composition is almost certain to be small. Students are in their twenty-fifth and twenty-sixth semesters of instruction in composition, having begun learning to write in the first grade. It is characteristic in situations in which instruction has been carried on over an extended period of time that the greatest change occurs early in the instruction, with the curve flattening out as instruction continues. With composition, rapid increases occur at the fifth and the seventh grade (cf. O'Donnell, p. 90). Though the O'Donnell study does not extend beyond the seventh grade, data in Hunt (p. 37) suggest that the mean improvement each year from the fourth to the twelfth grade is about 5 percent. Braddock (p. 7) gives the same figure.

The test for statistical significance is an estimate of the probability that the difference which is observed is so small as to be attributable to chance, or so large as to be attributable to instruction, or instruction plus maturation. Significance at the 5 percent level says that the difference observed would be likely to occur by chance only five times in one hundred if the given research were replicated. In working with probabilities, the chance that a given mean difference will be significant is greater for a large sample than for a small one. As may be seen from the discussion in this report (see discussion of first COOP table, p. 54), the chances of attaining a significant difference increase dramatically with an increase in the size of the experimental sample.

Thus when an investigator anticipates a small difference at best, a large sample is almost mandatory. In the opinion of the investigators, at least 100 pairs should be involved. It was partly to attain a larger sample that the investigators included other universities in the present study.

Matched Pairs

Particularly if an investigator is to follow college students beyond a single semester, the matched pairs design has the advantage of assuring continued comparability of the experimental and control subgroups. Students withdraw from college for many different reasons. The greater the number of academic terms (quarters or semesters) over which an experiment runs, the greater the number of students who will leave. The matched pairs design assures that such departures will not result in disparate subgroups. Particularly in regard to the importance of male-female differences in composition performance (see pp.167-168) differing proportions of men and women in the subgroups studied could easily distort results.

Another advantage of the matched pairs design (if students are matched exactly on one of the evaluative instruments) is that it permits a check on some of the computer programs developed to produce the statistical summaries and analyses. For example, in the present study an error on one of the computer runs occurred in the calculation of the t-ratios, used to test the significance of the difference between the experimental and control subgroup means. Had the pairs not been matched exactly on theme score, the error might have gone unnoticed. Similarly, the correlation coefficients between scores earned on two administrations of one of the objective tests were suspiciously low. Inspection of the data revealed that some of the scores had been aligned erroneously--Student A's score was assigned to Student B--and this had produced the low correlation coefficient. The ease with which such internal checks may be made certainly recommends the matched pairs design. Both of these checks on the accuracy of the data would have been difficult or impossible in a covariance design; both were easily made with the matched pairs design.

A further characteristic of the matched pairs design is that it permits the calculation of correlations between the subgroups. These across-subgroup correlations are, of course, essential for some of the analyses.

Thus, though the matched pairs design increases the rate of attrition, the investigators feel that the increased ease in making comparisons, in maintaining the similarity of the groups on the matching characteristics, and in detecting errors more than compensates for the difficulties. A more extensive presentation of the considerations which led to the employment of the matched pairs design is in Appendix B.

Themes as Tests

Though themes are properly used as evaluative instruments in research in composition, their use creates problems. Braddock (1963, pp. 6-15) discusses these problems at length and makes recommendations concerning ways of attempting to meliorate them. The present investigators, now completing their third treatments study, also have recommendations.

The central problem in using themes as evaluative instruments is that they do not lend themselves either to useful quantification or to consistent evaluation. To quantify demerits--for misspelling, for poor reference, for anemic development--will provide numbers which may then be manipulated, thus giving an impression of certainty and objectivity. But it is an illusion. The fact is that themes are primarily aesthetic objects, and judgments concerning them are aesthetic judgments. Each theme is unique; each judgment is to a considerable degree an expression of personal preference. Though themes may nonetheless be used in treatments research, the uniqueness of each theme and the large degree of subjectivity in the rating, force great care in evaluating procedures. In terms of these and other considerations, we make the following recommendations:

1. Students, or groups, should be matched exactly on sex. The present study demonstrates that among college freshmen, females as a group score higher on both objective and subjective tests of writing than do males. Failure to match on sex may easily lead to erroneous conclusions.

2. Since themes are unique aesthetic objects, they are influenced to some degree by the conditions under which they are produced. Performances on different testing occasions--such as at the beginning and at the end of a semester--should not be compared. Gain (change) scores are likely to be more misleading than enlightening. Rather, the students should be matched exactly on the evaluation of their initial performance, and compared on the second performance.

   Thus the second recommendation is that subgroups be matched on theme performance at the beginning of the study, and that the effectiveness of the treatment be assessed in terms of performance at the conclusion of the study.

3. The third recommendation is that only one topic be used for each test theme. Though it is true that some students will not do their best on the topic, it is also true that others will. The only way to provide more than one performance is to provide a different topic, and since the investigators believe each topic is unique, it seems better to assume that, _as a group_, the performance on one theme would be the same as the performance, _as a group_, on a second theme.* A corollary of this recommendation is that readers should evaluate only one topic at a time.

4. In treatments studies within a university, readers should follow an analytic procedure based upon the aspects or elements of composition stressed in the course. The determination of the efficacy of each treatment in producing the desired writing behavior is the goal of treatments research in composition. "The desired writing behavior" needs to be clearly defined and understood by the evaluators.

5. Testing conditions should be the same for both the experimental and the control subgroups. Not similar; the same. The best arrangement is for all participants in the study to write at the same time in the same room.

6. In a general investigation, such as the present one, in which the question is whether two samples of students drawn from five universities with five different composition programs perform differently when one sample has received instruction and the other has not, the theme evaluation _must_ be general--wholistic--rather than analytical. (Recommendation 4 refers to a specific course in a single institution, in which different procedures for attaining the same specific goals are being investigated.)

---

*Braddock advocates the use of two themes on each testing occasion, the better performance for each student to be used in the comparison, and a choice of topics. When the University of Iowa accepted the invitation to become a part of this study, Dr. Braddock requested and was granted permission to apply to the USOE for a separate grant. The grant was forthcoming. Using themes written for the present study in September 1964 and May 1966, together with separate themes written for him, by the same students, on each of those dates, Braddock made a comparison using the better theme by each student on each date. The study, _Evaluation of College-Level Instruction in Freshman Composition: Part II_, is complete and may be obtained by writing to Richard Braddock, Rhetoric Program, University of Iowa.

RECOMMENDATIONS

A characteristic of our research which was both a strength and a complication was the presence of multiple criteria. Three tests were employed, and these were administered at three different junctures in the students' first two college years. Furthermore, data were analyzed for both a constant N over the three testing occasions and for the maximum N on each testing occasion. The results, therefore, are not represented by a single, quantitative index. Instead, there are 18 sub-comparisons. The findings are not consistent among these. An inevitable characteristic of longitudinal research is some attrition of sample members. It was beyond the scope of the investigation to study directly the participants who dropped out at successive stages. Information such as overall scholastic averages, majors, and grades in specific courses might or might not have been useful in harmonizing the findings in the 18 comparisons.

With full realization of the complexities and the difficulty of arriving at a definitive interpretation of the evidence, the investigators offer some rather definite recommendations.

1. The investigators do not recommend the elimination of freshman English composition at this time.

   Data from this study suggest that required freshman composition as it was taught in the participating state universities during the period of this study had a definite effect on performance of the students tested at the end of the first semester, a less definite effect at the end of the second semester, and no effect at the end of the fourth semester. Because there was some evidence of superiority favoring those with composition instruction at two testing periods, the investigators do not recommend the elimination of freshman composition.

2. The investigators recommend that if the course is continued as a requirement, innovative practices be tried and their value assessed.

   The data do not strongly support the types of composition programs studied in this report; the investigators recommend further studies exploring the results of instruction centered on the new rhetorics, the new grammars, the production of films as stimulants to writing, small group instruction, individual instruction, speaking as a base for writing, and similar techniques which have been developed since the inception of the present study, to see if such approaches might be effective.

3. Course objectives in freshman composition should be stated in the most specific terms possible.

   "Improvement in writing" is a vague goal to set for the freshman course. It is particularly vague in view of the fact that the amount of improvement which may be expected, in the twenty-fifth and twenty-sixth semesters of the students' exposure to some kind of instruction in writing, is very small. In such a situation, one must specify "improvement" very carefully.

4. The present study contains basic information on test performance for a group of students who had proceeded through one, two, and four semesters of college without direct instruction in freshman composition. These data constitute a bench mark against which the performance of other groups can be compared. The investigators recommend such use.

5. The investigators recommend that institutional norms and national norms for tests designed to measure performance in writing be set up for males and females separately. Results of research which does not separate male and female performance should be interpreted with care.

REFERENCES

1. Braddock, Richard, Richard Lloyd-Jones, and Lowell Schoer. *Research in Written Composition*. Champaign, Illinois: National Council of Teachers of English, 1963.

2. Braddock, Richard, and Charles R. Statler. *Evaluation of College-Level Instruction in Freshman Composition: Part II* (Cooperative Research Project No. S-260). Iowa City, Iowa: University of Iowa, April, 1968.

3. "Composition Test Shows High Validity on Reliable Criterion of Writing Ability," *ETS Developments*, XI (January 1963), 1 & 4.

4. Godschalk, Fred, Frances Swineford, and William E. Coffman. *The Measurement of Writing Ability*. New York: College Entrance Examination Board, 1966.

5. Hunt, Kellogg W. *Grammatical Structures Written at Three Grade Levels*. National Council of Teachers of English, Research Report No. 3. Chicago, Illinois: National Council of Teachers of English, 1965.

6. Jewell, Ross M., *et al*. *Final Report of the Communication Experiment Conducted by the Department of Languages, Speech, and Literature of the Iowa State Teachers College, 1955-58*. Cedar Falls, Iowa: Iowa State Teachers College, 1960.

7. Jewell, Ross M., and Gordon J. Rhum. *The Relative Effectiveness of Two Methods of Instruction in College Freshman Composition: Closed-Circuit Television and 'Normal' Classroom*. Cedar Falls, Iowa: State College of Iowa, February, 1966.

8. Jewell, Ross, John Cowley, and Gordon Rhum. *Interim Report: The Effectiveness of College-Level Instruction in Freshman Composition* (Cooperative Research Project 2188). Cedar Falls, Iowa: State College of Iowa, 1966.

9. Loban, Walter. *Language Ability Grades Seven, Eight, and Nine* (Cooperative Research Monograph No. 18). Washington, D. C.: Government Printing Office, 1966.

10. O'Donnell, Roy C., William J. Griffin, Raymond C. Norris. *Syntax of Kindergarten and Elementary School Children: A Transformational Analysis*. National Council of Teachers of English, Research Report No. 8. Chicago, Illinois: National Council of Teachers of English, 1967.

11. Riling, Mildred E. *Oral and Written Language of Children in Grades 4 and 6 Compared with the Language of Their Textbooks* (Cooperative Research Project 2410). Durant, Oklahoma: Southeastern State College of Oklahoma, 1965.

12. Roberts, Holland. [A review of the CEEB English Composition Test], *Sixth Mental Measurements Yearbook*. Ed. Oscar K. Buros. Highland Park, New Jersey: Gryphon Press, 1965, pp. 589-91.

13. Rulon, Phillip J., "A Simplified Procedure for Determining the Reliability of a Test by Split-Halves," *Harvard Educational Review*, IX (1939), 99-103.

APPENDIX A

Theme Topics and Instructions

APPENDIX A

Theme Topics and Instructions

The principles followed in selecting topics, the use of a single topic on each test administration, and the equivalence of topics actually employed need to be discussed briefly.

Three criteria were established in selecting topics for the theme tests: the topic must be of a middle level of abstraction, it must be related to the students' experience, and it must call for an individual rather than a stock response. A middle level of abstraction avoided favoring either the students who were skillful in exploring general principles or the students who happened to have special knowledge related to a specific topic. A topic related to the students' experience and knowledge allowed them to support and illustrate their general statements with particulars readily available to them. A topic calling for an individual rather than a stock response provided a test of the students' ability to establish and support an original thesis.

The use of a single topic rather than a choice among several topics on each testing occasion avoided the introduction of an additional variable whose influence would be difficult to estimate. Such a restriction seemed justified by the fact that the students' performance as individuals was not under investigation. There is no reason to believe that if the students had had a choice of topics, comparison of their group performance would have been different from that resulting from a single topic.

Equivalence of topics across testing occasions was not vital, as students' change scores on theme performance were not considered in the conclusions in this study. Though it was hoped that the topics used would be comparable to one another, any lack of similarity which may be present cannot be used meaningfully in speculation about the results achieved. The subgroups were compared with one another on their performance at each testing occasion. Changes from occasion to occasion within subgroups were not investigated.

On the following pages are the instructions and theme topics for the various testing sessions. The complete instruction sheets, with places for the readers' ratings, the name and number of the student, and the like have not been reproduced as these details are irrelevant and reproduction difficult. It should be noted, however, that the original instruction sheets were so arranged that the graders could learn neither the student's name nor the date on which the paper was written, and the second reader could not see the rating given the paper by the first reader.

(Theme Instructions for September 1964)

THEME INSTRUCTIONS

1. The paper which you are about to write will be judged on your success in presenting your thoughts in a clear, unified, well-organized manner, observing the conventions of standard written English. You should think about the topic until you have determined what idea you want to convey to the reader and the general procedure you will follow in doing so. Then you may write your paper. Do not hesitate to make a brief outline *if you desire to do so* (use the back of this sheet). An outline is *not* required.

2. You should write as neatly and legibly as you can, but you should not hesitate to make changes between the lines if you believe them to be necessary. You do not have to copy the paper over.

3. WRITE ON ONE SIDE OF THE PAPER ONLY. If you need more paper, ask for it.

4. Begin on the third line of the first sheet, and WRITE ON EVERY LINE THEREAFTER.

5. You must write with INK or BALL-POINT PEN.

6. Be certain to write your STUDENT NUMBER in each of the blanks (two at the top, one at the bottom) provided for it on this sheet, and in the upper right-hand corner of *each page* of your theme.

7. Turn in all of the paper given to you.

8. You must stay at least one hour and fifteen minutes.

9. LENGTH: 300 - 500 words.

TOPIC

Today a young man who wears a beard or a girl who prefers slacks to skirts has difficulty in finding employment in most work which serves the public. Changes in fashion are announced one day and adopted the next. In business, promotions are made with great emphasis upon how well an individual meets the "image" the employer wishes to create. In school, those who do as they are told and give the answers expected of them are rated high by many of the faculty; those who do what "everyone else" does are popular with the students.

Now consider a famous quotation: "Whoso would be a man must be a non-conformist."

Relate the material in the opening paragraph to the quotation, indicating whether, on the basis of your observation and experience, you feel the idea expressed in the quotation is true.

(Theme Instructions for January 1965)

THEME INSTRUCTIONS

1. The paper which you are about to write will be judged on your success in presenting your thoughts in a clear, unified, well-organized manner, observing the conventions of standard written English. You should think about the topic until you have determined what idea you want to convey to the reader and the general procedure you will follow in doing so. Then you may write your paper. Do not hesitate to make a brief outline if you desire to do so (use the back of this sheet). An outline is not required.

2. You should write as neatly and legibly as you can, but you should not hesitate to make changes between the lines if you believe them to be necessary. You do not have to copy the paper over.

3. WRITE ON ONE SIDE OF THE PAPER ONLY. If you need more paper, ask for it.

4. Begin on the third line of the first sheet, and WRITE ON EVERY LINE THEREAFTER.

5. You must write with INK or BALL-POINT PEN.

6. Be certain to write your STUDENT NUMBER in the blank provided at the top of this instruction sheet in the upper left-hand corner under the Total Score box. It should also be written on each page of your theme. Do NOT write your name, or the name of your school, in any place other than the blank provided at the bottom of this sheet.

7. Turn in all of the paper given to you.

8. You must stay at least one hour and fifteen minutes.

9. LENGTH: 300 - 500 words.

TOPIC

In the United States, popular entertainment reflects the ideals of the great middle class of people. For example, we seldom see or read of a young couple struggling to make ends meet, of psychological problems that cannot be resolved, of the blood that accompanies violent death, of the horrors of war, or of the wearing routine of life day in and day out. On the contrary, no problem is too complex for solution, no disaster occurs to the Good, no reward to the Bad.

It is hardly too much to say that most young people in the United States form their expectation of their lives as adults from the distorted image presented by television, movies, and books rather than from their observations of the lives of the adults about them.

Reflect upon these statements and determine whether you agree or disagree with them or feel that they should be modified in some way. Then write a paper indicating the manner in which your experience and knowledge have led you to the conclusion you have reached.

(Theme Instructions for May 1965)

THEME INSTRUCTIONS

1. The paper which you are about to write will be judged on your success in presenting your thoughts in a clear, unified, well-organized manner, observing the conventions of standard written English. You should think about the topic until you have determined what idea you want to convey to the reader and the general procedure you will follow in doing so. Then you may write your paper. Do not hesitate to make a brief outline _if you desire to do so_ (use the back of this sheet). An outline is _not_ required.

2. You should write as neatly and legibly as you can, but you should not hesitate to make changes between the lines if you believe them to be necessary. You do not have to copy the paper over.

3. WRITE ON ONE SIDE OF THE PAPER ONLY. If you need more paper, ask for it.

4. Begin on the third line of the first sheet, and WRITE ON EVERY LINE THEREAFTER.

5. You must write with INK or BALL-POINT PEN.

6. Be certain to write your STUDENT NUMBER in the blank provided at the top of this instruction sheet in the upper left-hand corner under the Total Score box. It should also be written on each page of your theme. _Do NOT write your name, or the name of your school, in any place other than the blank provided at the bottom of this sheet._

7. Turn in all of the paper given to you.

8. You must stay at least one hour and fifteen minutes.

9. LENGTH: 300 - 500 words.

TOPIC

As society becomes increasingly complex, the number of people upon whom we are dependent increases. Daniel Boone killed a bear and ate it. When we buy steak, we purchase the services of the person who produced the animal, the person who fattened it, the person who took it to market, the packing company which bought it, slaughtered it, and dressed it, the trucker who transported it to the store from which we bought it, and, of course, the grocer himself. Each person must do his part if we are to have the steak. Even this picture is greatly over-simplified. There are, for example, the gasoline which fueled the truck and the truck itself. Considering the interdependence illustrated by the story of the steak, how free are we to guide our own lives? Are we liberated from stalking, killing, skinning, and cleaning our dinner, or are we robbed of our independence? Can we say, as Henley did, "I am the master of my fate, /I am the captain of my soul"? Does modern technology liberate us or dominate us? Present your opinion, based upon your knowledge, observation, and experience.

(Theme Instructions for May 1966)

THEME INSTRUCTIONS

1. The paper which you are about to write will be judged on your success in presenting your thoughts in a clear, unified, well-organized manner, observing the conventions of standard written English. You should think about the topic until you have determined what idea you want to convey to the reader and the general procedure you will follow in doing so. Then you may write your paper. Do not hesitate to make a brief outline _if you desire to do so_ (use the back of this sheet). An outline is _not_ required.

2. You should write as neatly and legibly as you can, but you should not hesitate to make changes between the lines if you believe them to be necessary. You do not have to copy the paper over.

3. WRITE ON ONE SIDE OF THE PAPER ONLY. If you need more paper, ask for it.

4. Begin on the third line of the first sheet, and WRITE ON EVERY LINE THEREAFTER.

5. You must write with INK or BALL-POINT PEN.

6. Be certain to write your STUDENT NUMBER in the blank provided at the top of this instruction sheet in the upper left-hand corner under the Total Score box. It should also be written on each page of your theme. _Do NOT write your name, or the name of your school, in any place other than the blank provided at the bottom of this sheet._

7. Turn in all of the paper given to you.

8. You must stay at least one hour and fifteen minutes.

9. LENGTH: 300 - 500 words.

TOPIC

Conventional is a word frequently used to refer to customary attitudes, beliefs or actions. In the United States it is a convention for men to be clean-shaven, women to wear a certain amount of make-up, boys to be interested in sports, and girls to be interested in becoming wives and mothers. A person who is unconventional in some way departs from the conventions of action or belief of the society of which he is a part.

With this explanation in mind, discuss the following statement:

"Convention is society's safeguard, but also its potential executioner." To what extent and in what ways do you agree with this statement? Use examples and details from your knowledge and experience to support your conclusion.

APPENDIX B

Choice of Experimental Design

APPENDIX B

Choice of Experimental Design

In planning research, some of the most complex questions are concerned with the choice of experimental design. The questions are both theoretical and functional. These two kinds of consideration come together when one finally must decide the best way, under the circumstances in which a given study will be made, to collect and analyze data for meaningful samples of students. In the present study, three circumstances dictated the choice of a matched pairs design.

The first circumstance was the college administration's stipulation that students who were to receive the experimental treatment be informed of the fact prior to their registration. It seemed essential that such students, their parents, and the faculty advisors be given advance information about the purpose of the research and its impact on them. These experimental students would not receive instruction in freshman composition--a major departure from normal college experience. Given the faith of students in the importance of composition,[1] to have denied them enrollment on registration day without prior warning could have induced anxiety and resentment, possibly producing a kind of "reverse" Hawthorne effect. Added to this would have been confusion in registration, irritation among advisors, and concern among parents.

Thus the investigators were compelled to select, in advance of September registration, the students who would receive the experimental treatment. As described on page 38, this procedure involved selecting a pool of students from those who, by approximately July 1, 1964, had met admission requirements and expressed their intention to enroll in the given institution. There was, of course, no assurance that all of the selected pool would actually enroll. This pool, which was a random sample from the July list, would not be a random sample of the September freshman class. That is, some entering freshman students had no opportunity to be included, and some who were included in the July group did not enroll.

A second circumstance was the duration of the investigation. The experimental design called for the students to be tested through the end of their sophomore year. That relatively heavy attrition would occur was certain;* that it would have an equal effect on both treatment groups seemed unlikely. Among other considerations, the control students would be enrolled in a course which frequently causes students trouble, while the experimental students would not. In any event, the possibility that attrition would occur in such a way that the two treatment groups would become progressively dissimilar could not be ignored.

---

[1] Jewell, Ross M. and Gordon J. Rhum, The Relative Effectiveness of Two Methods of Instruction in College Freshman Composition: Closed-Circuit Television and 'Normal' Classroom. Cedar Falls, Iowa: State College of Iowa, February, 1966, p. 48.

Related to the attrition problem was the importance of maintaining the same ratio of males to females in both of the subgroups. The investigators believed, and their belief is supported by data subsequently examined (see page 140),that females would perform somewhat better than males on measures of composition ability. Should the ratio between sexes in one group become substantially different from the ratio in the other group, the likelihood of distorted results would be present.

A third circumstance was the audience which would read the research. As the investigation concerns the effectiveness of a course usually taught in departments of English, members of English departments would be the group for whom the report was primarily intended. It seems fair to say that such an audience would have considerable difficulty in following the intricacies of analysis of covariance. Though this consideration may at first seem somewhat frivolous, its pertinence to the potential impact of the project is nonetheless real.

In the light of these circumstances, the investigators became convinced that the matched-pairs design should be employed. Matching after September registration insured a list of students who were actually enrolled. Use of the matched pairs design with sex as one criterion made certain that the ratio between males and females would be the same for both subgroups not only at the beginning, but at any subsequent point in the investigation. Use of matched pairs minimized the possibility that in the attrition which would occur over the life of the experiment some factor would operate unequally to reduce the similarity of the subgroups. Finally, use of matched pairs enabled the investigators to present results in a manner which would make them readily available to members of English departments and directors of freshman composition.

---

*The Registrar of the University of Northern Iowa estimates that the attrition for a freshman class is on the order of 19 percent, and the attrition has reached approximately 40 percent by the end of the sophomore year.

The investigators could, of course, have set up the subgroups from among the students whose data were available in July, taking first a random sample of the total group, pairing them, and then for each matched pair of students randomly assigning one member of the pair to the experimental treatment and the other member of the pair to the control treatment. However, in July the only pertinent test data available for the students was their performance on ACT English. As the investigators wished to match as closely as possible, they decided to wait until more tests could be administered during the fall semester orientation period. Doing so permitted matching as reported on page 39, by age, sex, theme performance, and a score derived from performance on the CEEB and COOP. This precision in matching provided increased confidence in the similarity between the two treatment groups. Closeness in matching was also facilitated by the fact that the supply of subjects was greater in September than it was in July.

Three additional points. Since there were only two treatment groups, the matched pairs approach was more feasible than if there had been several treatment groups. Secondly, the investigators did not have to use, indeed did not wish to use, intact classroom groups for the control treatment. Finally, in methods experiments generally, random samples of a real population are not attainable. Near-randomness is achieved only in the beginning stages, and not in the groups which actually complete the experimental period.

APPENDIX C

Procedure for Evaluating Themes

APPENDIX C

Procedure for Evaluating Themes

Prior to each reading session, Mr. Jewell would send to Mr. Godshalk about forty themes, selected at random. From this sample Mr. Godshalk would determine the general nature of the total set of themes. He would choose a number of themes that in his judgment were typical of range and treatment, and Mr. Jewell would have these duplicated. These became the sample themes used during the reading as practice themes.

Mr. Godshalk's main responsibility when the raters (the smallest number was nine) had assembled was to communicate to them the criteria for evaluating the papers. First, he would have Mr. Cowley and Mr. Jewell describe the purpose of the investigation, the circumstances under which the papers had been written, and the students who had written them. He would then explain the rating scale. When all questions concerning its application had been answered, he would distribute several sample themes to be rated. After he had made a tally of the various values assigned to these papers, he would allow individuals to explain their ratings or to question his rating. If a rater seemed to be over-reacting to something in the papers, something which Mr. Godshalk believed from examination of the sample papers was typical, he would so inform the readers and caution them against misinterpreting particular aspects of the papers.

Before setting the readers to work in earnest, he would remind them that since they were experienced readers their first judgment of a theme as a whole was probably as valid as any subsequent judgment they might make of the same paper. Therefore, they were not to pause and consider but were to read and respond. As the rating session progressed, Mr. Godshalk would note whether any particular rater seemed to judge consistently in a way different from the other raters. At relatively frequent intervals, he would interrupt the reading to allow the readers to relax and would read aloud papers which had been passed on to him by individual readers. Frequently, these papers posed special problems which Mr. Godshalk would have the group discuss, always making clear his own judgment. The goal of the initial orientation and of the subsequent breaks in the reading was for Mr. Godshalk to convey to the readers his criteria and to get them to standardize their scoring so that they would agree in their ratings. The reading would be most "perfect" when all of the readers rated all the papers in the same way that Mr. Godshalk would rate them. In practice his standards would be slightly altered if a consensus indicated they should be. Thus, the validity of the evaluation could be no greater than the validity of Mr. Godshalk's criteria as modified on occasion by discussion with the readers.

APPENDIX D

Performance in Individual Universities

CONTENTS

Page

Performance in Individual Universities 179
  First Semester 179
    COOP 179
    CEEB 179
    Theme 180
  Second Semester 180
    COOP 180
    CEEB 180
    Theme 180
  First Year 181
    COOP 181
    CEEB 181
    Theme 181

LIST OF TABLES
All 5 Universities

Table		Page
D-I	The Performance of 597 Matched Pairs of Students on the Cooperative English Tests: English Expression at the End of the First Semester (January 1965) by Individual Universities	182
D-II	The Performance of 597 Matched Pairs of Students on the College Entrance Examination Board English Composition Test at the End of the First Semester (January 1965) by Individual Universities	183
D-III	The Performance of 597 Matched Pairs of Students on Total Theme Rating at the End of the First Semester (January 1965) by Individual Universities	184
D-IV	The Performance of 365 Matched Pairs of Students on the Cooperative English Tests: English Expression at the End of the Second Semester (May 1965) by Individual Universities	185
D-V	The Performance of 365 Matched Pairs of Students on the College Entrance Examination Board English Composition Test at the End of the Second Semester (May 1965) by Individual Universities	186
D-VI	The Performance of 365 Matched Pairs of Students on Total Theme Rating at the End of the Second Semester (May 1965) by Individual Universities	187
D-VII	The Performance of 365 Matched Pairs of Students on the Cooperative English Tests: English Expression at the End of the First Academic Year (May 1965) by Individual Universities	188
D-VIII	The Performance of 365 Matched Pairs of Students on the College Entrance Examination Board English Composition Test at the End of the First Academic Year (May 1965) by Individual Universities	189
D-IX	The Performance of 365 Matched Pairs of Students on Total Theme Rating at the End of the First Academic Year (May 1965) by Individual Universities	190

PERFORMANCE IN INDIVIDUAL UNIVERSITIES

The project was designed to yield information regarding instruction in college freshman composition at state-supported universities offering varied freshman composition programs. In the light of the broad purpose of the study, comparison of performance among individual universities was neither a primary nor a secondary objective. That is, no attempt has been made to assess the apparent effectiveness or lack of effectiveness of the programs at the individual universities. Rather, the focus has been on the total group of students.

The investigators agreed that each university should receive a report of the results relating to its own students. Presented on pages 182 to 190 are summary tables of some of the basic facts of the performance at individual universities. Interpretation of the evidence in these tables must be tentative, primarily because the samples are so small. Also because of small samples, no summary of May 1965-May 1966 performance is included.

Below are summary statements based on Tables D-I through D-IX. In preparing the following summary statements the investigators identified only what seemed to be the most prominent departures from the composite picture for the participating universities.

First Semester (September 1964-January 1965)

COOP. (September to January gains)

1. Gains for university 1 were in general greater than those for any other university.

2. For university 5, gains made by the controls, both males and females, were relatively low.

CEEB. (September to January gains)

1. The gains at university 1 were in general greater than those for the combined universities.

2. In universities 2 and 3, the mean gains were generally below those for the combined universities.

3. University 5 is special, as it was the only one at which the experimentals ended the semester higher than the controls.

Theme. (January 1965 means only, no gain scores for themes analyzed)

1. In university 1, the control minus experimental mean difference was greater than it was for the combined universities.
2. In university 2, the experimentals performed somewhat better than the controls.
3. In university 3 the mean theme scores were lower than the mean for the combined universities. This was especially true of the males.

Second Semester (January 1965-May 1965)

COOP. (January to May gains)

1. University 5 had the greatest gains.
2. University 2 had the smallest gains.
3. In university 1 the control minus experimental value was greatest.
4. Among the males, the mean gains in university 1 were smallest.

CEEB. (January to May gains)

1. In university 1 the control minus experimental difference was relatively large.
2. In university 5 at the end of the freshman year, the mean for experimentals was greater than the mean for controls.
3. Within universities there was fluctuation in mean gains between experimentals and controls, males and females.

Theme. (May means only, no gain scores for themes analyzed)

1. University 2 had mean theme scores higher than those for the combined universities.

2. University 3 had mean theme scores lower than those for the combined universities.

3. Difference in theme score in favor of controls was greatest at university 3.

First Year (September 1964-May 1965)

COOP. (September to May gains)

1. The greatest gain was made by the experimental males at university 5.

2. The greatest control minus experimental difference at the end of the second semester was at university 1.

CEEB. (September to May gains)

1. The highest gains were made by the female controls at university 1.

2. The lowest gains were made by female experimentals at university 2.

3. The greatest control minus experimental mean difference in May 1965 was at university 1.

Theme. Summary above under Second Semester.

TABLE D-I

THE PERFORMANCE OF 597 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION AT THE END OF THE FIRST SEMESTER (JANUARY 1965) BY INDIVIDUAL UNIVERSITIES

				Total Group				Experimentals				Controls				
				Experimentals		Controls		Males		Females		Males		Females		Jan. Diff.
Univ.	N	M	F	Jan. Mean	Gain	Jan. Mean	Gain	Jan. Mean	Gain	Jan. Mean	Gain	Jan. Mean	Gain	Jan. Mean	Gain	Controls minus Exp.
All	597	235	362	163.93	1.64	164.94	2.70	161.64	1.64	165.41	1.64	162.70	2.84	166.39	2.61	1.01*
1	200	65	135	163.42	2.37	164.56	3.78	161.43	3.94	164.37	1.61	161.34	3.83	166.11	3.75	1.14*
2	127	65	62	165.67	2.10	166.50	2.87	164.89	2.37	166.48	1.82	165.86	3.40	167.16	2.32	0.83
3	124	34	90	164.59	1.19	165.57	2.81	159.56	0.74	166.49	1.36	160.62	3.88	167.44	2.41	0.98
4	20	9	11	156.05	-4.75	161.55	1.00	145.56	-13.22	164.64	2.18	161.11	3.78	161.91	-1.27	5.50
5	126	62	64	163.58	1.46	163.89	0.98	161.92	1.10	165.19	1.81	162.19	0.50	165.53	1.44	0.31

*Significant at 0.05 level (two-tailed test).

TABLE D-II

THE PERFORMANCE OF 597 MATCHED PAIRS OF STUDENTS ON THE COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TEST AT THE END OF THE FIRST SEMESTER (JANUARY 1965) BY INDIVIDUAL UNIVERSITIES

				Total Group				Experimentals				Controls				
				Experimentals		Controls		Males		Females		Males		Females		
Univ.	N	M	F	Jan. Mean	Gain	Jan. Mean	Gain	Jan. Mean	Gain	Jan. Mean	Gain	Jan. Mean	Gain	Jan. Mean	Gain	Jan. Diff. Controls minus Exp.
All	597	235	362	507.12	29.37	508.88	30.47	486.45	32.53	520.54	27.32	488.91	34.32	521.83	27.98	1.76
1	200	65	135	500.37	43.38	507.05	46.15	472.35	47.08	513.86	41.60	479.49	54.28	520.32	42.24	6.68
2	127	65	62	508.49	11.57	511.73	15.61	495.75	7.83	521.84	15.48	509.97	22.91	513.58	7.95	3.24
3	124	34	90	502.33	18.60	507.88	18.35	458.03	33.79	519.07	12.86	474.12	28.82	520.63	14.39	5.55
4	20	9	11	497.70	26.95	495.85	21.55	479.22	31.11	512.82	23.55	491.00	29.56	499.82	15.00	-1.85
5	126	62	64	522.66	36.07	511.94	33.91	508.10	42.68	536.77	29.67	484.53	29.06	538.50	38.61	-10.71

TABLE D-III

THE PERFORMANCE OF 597 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING AT THE END OF THE FIRST SEMESTER (JANUARY 1965) BY INDIVIDUAL UNIVERSITIES

				Total Group		Experimentals		Controls		
Univ.	N	M	F	Experimentals Jan. Mean	Controls Jan. Mean	Males Jan. Mean	Females Jan. Mean	Males Jan. Mean	Females Jan. Mean	Jan. Diff. Controls minus Exp.
All	597	235	362	9.69	9.96	8.77	10.29	9.22	10.44	0.27*
1	200	65	135	9.50	10.26	8.18	10.14	9.28	10.73	0.75*
2	127	65	62	10.31	10.04	9.75	10.90	9.57	10.53	-0.28
3	124	34	90	9.10	9.34	7.50	9.71	7.71	9.96	0.23
4	20	9	11	9.80	10.05	8.22	11.09	9.56	10.45	0.25
5	126	62	64	9.91	10.02	9.13	10.67	9.58	10.44	0.10

*Significant at 0.05 level (two-tailed test).

TABLE D-IV

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION AT THE END OF THE SECOND SEMESTER (MAY 1965) BY INDIVIDUAL UNIVERSITIES

				Total Group				Experimentals				Controls				
				Experimentals		Controls		Males		Females		Males		Females		May Diff.
Univ.	N	M	F	May Mean	Gain	May Mean	Gain	May Mean	Gain	May Mean	Gain	May Mean	Gain	May Mean	Gain	Controls minus Exp.
All	365	134	231	166.50	1.99	167.29	1.53	164.63	1.87	167.58	2.06	164.91	1.60	168.68	1.49	0.79*
1	137	34	103	165.16	1.47	166.53	0.85	160.59	-0.50	166.67	2.12	162.41	0.18	167.88	1.07	1.36*
2	98	42	56	167.85	0.43	168.27	0.13	166.95	0.57	168.52	0.32	166.67	-0.07	169.46	0.29	0.42
3	40	16	24	166.15	3.67	166.23	1.65	162.00	3.44	168.92	3.83	162.50	1.50	168.71	1.75	0.07
5	90	42	48	167.22	3.72	167.88	4.03	166.60	4.48	167.77	3.06	166.10	4.45	169.44	3.67	0.66

*Significant at 0.05 level (two-tailed test).

NOTE: Gains are for one semester, January 1965–May 1965 only. After the first semester the data at university 4 were limited and incomplete and therefore not included in the totals.

TABLE D-V

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON THE COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TEST AT THE END OF THE SECOND SEMESTER (MAY 1965) BY INDIVIDUAL UNIVERSITIES

				Total Group				Experimentals				Controls				
				Experimentals		Controls		Males		Females		Males		Females		May Diff.
Univ.	N	M	F	May Mean	Gain	May Mean	Gain	May Mean	Gain	May Mean	Gain	May Mean	Gain	May Mean	Gain	Controls Minus Exp.
All	365	134	231	516.98	2.65	524.28	12.74	501.47	7.01	525.98	0.13	500.28	13.40	538.20	12.36	7.30
1	137	34	103	508.61	-1.63	527.71	16.89	478.24	-8.18	518.63	0.53	482.88	3.65	542.50	21.26	19.10*
2	98	42	56	528.17	9.53	537.06	11.04	523.95	18.71	531.34	2.64	526.45	13.81	545.02	8.97	8.89
3	40	16	24	497.85	3.37	495.55	10.90	473.50	30.87	514.08	-14.96	476.31	12.56	508.38	9.79	-2.30
5	90	42	48	526.06	1.36	517.91	9.10	508.45	-1.50	541.46	3.85	497.31	21.19	535.94	-1.48	-8.14

*Significant at 0.05 level (two-tailed test).

NOTE: Gains are for one semester, January 1965-May 1965 only. After the first semester the data at university 4 were limited and incomplete and therefore not included in the totals.

TABLE D-VI

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING AT THE END OF THE SECOND SEMESTER (MAY 1965) BY INDIVIDUAL UNIVERSITIES

				Total Group		Experimentals		Controls		
Univ.	N	M	F	Experimentals May Mean	Controls May Mean	Males May Mean	Females May Mean	Males May Mean	Females May Mean	May Diff. Controls minus Exp.
All	365	134	231	9.63	9.76	9.30	9.82	9.39	9.97	0.13
1	137	34	103	9.36	9.43	8.41	9.67	8.91	9.60	0.07
2	98	42	56	10.35	10.39	10.50	10.23	9.88	10.77	0.04
3	40	16	24	8.40	8.88	7.81	8.79	8.44	9.17	0.48
5	90	42	48	9.81	9.96	9.38	10.19	9.64	10.23	0.14

NOTE: After the first semester the data at university 4 were limited and incomplete and therefore not included in the totals.

TABLE D-VII

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION AT THE END OF THE FIRST ACADEMIC YEAR (MAY 1965) BY INDIVIDUAL UNIVERSITIES

				Total Group				Experimentals				Controls				
				Experimentals		Controls		Males		Females		Males		Females		May Diff.
Univ.	N	M	F	May Mean	Gain	May Mean	Gain	May Mean	Gain	May Mean	Gain	May Mean	Gain	May Mean	Gain	Controls minus Exp.
All	365	134	231	166.50	3.90	167.29	4.50	164.63	4.13	167.58	3.77	164.91	4.28	168.68	4.63	0.79*
1	137	34	103	165.16	3.46	166.53	5.12	160.59	2.94	166.67	3.63	162.41	4.91	167.88	5.19	1.36*
2	98	42	56	167.85	3.32	168.27	3.44	166.95	3.62	168.52	3.09	166.67	3.45	169.46	3.43	0.42
3	40	16	24	166.15	3.80	166.23	4.50	162.00	2.87	168.92	4.42	162.50	5.94	168.71	3.54	0.07
5	90	42	48	167.22	5.24	167.88	4.71	166.60	6.07	167.77	4.52	166.10	3.95	169.44	5.37	0.66

*Significant at 0.05 level (two-tailed test).

NOTE: Gains are for the first full academic year, September 1964-May 1965.

TABLE D-VIII

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON THE COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TEST AT THE END OF THE FIRST ACADEMIC YEAR (MAY 1965) BY INDIVIDUAL UNIVERSITIES

				Total Group				Experimentals				Controls				
				Experimentals		Controls		Males		Females		Males		Females		May Diff. Controls
Univ.	N	M	F	May Mean	Gain	May Mean	Gain	May Mean	Gain	May Mean	Gain	May Mean	Gain	May Mean	Gain	minus Exp.
All	365	134	231	516.98	32.79	524.28	41.47	501.47	33.98	525.98	32.11	500.28	33.92	538.20	45.84	7.30
1	137	34	103	508.61	44.66	527.71	59.18	478.24	48.35	518.63	43.45	482.88	50.82	542.50	61.94	19.10*
2	98	42	56	528.17	12.05	537.06	24.00	523.95	16.16	531.34	8.96	526.45	17.24	545.02	29.07	8.89
3	40	16	24	497.85	28.85	495.55	20.60	473.50	41.87	514.08	20.17	476.31	19.25	508.38	21.50	-2.30
5	90	42	48	526.06	39.07	517.91	42.79	508.45	37.14	541.46	40.75	497.31	42.50	535.94	43.04	-8.14

*Significant at 0.05 level (two-tailed test).

NOTE: Gains are for the first full academic year, September 1964-May 1965.

TABLE D-IX

THE PERFORMANCE OF 365 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING AT THE END OF THE FIRST ACADEMIC YEAR (MAY 1965) BY INDIVIDUAL UNIVERSITIES

				Total Group		Experimentals		Controls		
Univ.	N	M	F	Experimentals May Mean	Controls May Mean	Males May Mean	Females May Mean	Males May Mean	Females May Mean	May Diff. Controls Minus Exp.
All	365	134	231	9.63	9.76	9.30	9.82	9.39	9.97	0.13
1	137	34	103	9.36	9.43	8.41	9.67	8.91	9.60	0.07
2	98	42	56	10.35	10.39	10.50	10.23	9.88	10.77	0.04
3	40	16	24	8.40	8.88	7.81	8.79	8.44	9.17	0.48
5	90	42	48	9.81	9.96	9.38	10.19	9.64	10.23	0.14

LIST OF TABLES
University 1

Table		Page
D-I(1)	Achievement as of September 1964 of the Sample of 1964-65 Freshman Students and of Various Persisting Portions of that Sample: University 1	195
D-II(1)	Performance of Available Samples of Matched Pairs of Students on Three Criterion Measures at Beginning, Middle, and End of First Year and End of Second Year of College: University 1	197
D-III(1)	The Performance of 200 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 1	198
D-IV(1)	The Performance of 200 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 1	199
D-V(1)	The Performance of 200 Matched Pairs of Students on Total Theme Rating in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 1	200
D-VI(1)	The Performance of 137 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 1	201
D-VII(1)	The Performance of 137 Matched Pairs of Students on the Cooperative English Tests: English Expression in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 1	202

Table | Page

D-VIII(1) The Performance of 137 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 1 203

D-IX(1) The Performance of 137 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 1 204

D-X(1) The Performance of 137 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 1 205

D-XI(1) The Performance of 137 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 1 206

D-XII(1) The Performance of 137 Matched Pairs of Students on Total Theme Rating in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 1 207

D-XIII(1) The Performance of 137 Matched Pairs of Students on Total Theme Rating in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 1 208

D-XIV(1) The Performance of 137 Matched Pairs of Students on Total Theme Rating in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 1 209

Table | Page

D-XV(1) The Performance of 43 Matched Pairs of Students on Cooperative English Tests: English Expression in September 1964, January 1965, May 1965, and May 1966; Including Differences in Means, Standard Deviations, and t-Ratios: University 1 210

D-XVI(1) The Performance of 43 Matched Pairs of Students on the College Entrance Examination Board English Composition Test in September 1964, January 1965, May 1965, and May 1966; Including Differences in Means, Standard Deviations, and t-Ratios: University 1 211

D-XVII(1) The Performance of 43 Matched Pairs of Students on the Total of Two Theme Ratings in September 1964, January 1965, May 1965, and May 1966; Including Differences in Means, Standard Deviations, and t-Ratios: University 1 212

TABLE D-I(1)

ACHIEVEMENT AS OF SEPTEMBER 1964 OF THE SAMPLE OF 1964-65 FRESHMAN STUDENTS AND OF VARIOUS PERSISTING PORTIONS OF THAT SAMPLE: UNIVERSITY 1

Sample and Subgroups	N	Percent Men	Percentile Rank in H. S. Class		ACT English Stan. Score		ACT Composite Stan. Score		COOP English Exp. Converted Score	
			Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.
Exp. Pool plus Control Pool	901 to 943	38.5	69.67	17.97	22.35	3.33	23.30	3.30	160.91	8.34
Exp. Pool	284 to 296	37.5	70.00	17.44	22.28	3.34	23.12	3.12	160.54	7.93
Control Pool	617 to 647	38.9	69.52	18.20	22.37	3.32	23.38	3.38	161.09	8.53
Matched Exp. September 1964	224	33.9	70.27	15.95	22.26	3.19	22.96	2.94	160.78	7.30
Matched Controls September 1964	224	33.9	71.34	17.22	22.50	3.19	23.41	2.99	160.49	7.36
Exp. Group January 1965	200	34.5	70.30	16.09	22.34	3.21	22.97	2.94	161.05	7.29
Control Group January 1965	200	34.5	72.25	17.04	22.56	3.15	23.48	2.90	160.79	7.32
Exp. Group May 1965	137	24.8	71.31	15.03	22.62	2.95	23.10	2.75	161.70	7.26
Control Group May 1965	137	24.8	73.21	17.75	22.69	3.01	23.44	2.74	161.40	7.53
Exp. Group May 1966	43	13.9	75.12	13.87	Not available				162.91	6.76

TABLE D-I(1)
CONTINUED

Sample and Subgroups	N	CEEB English Composition Stan. Rating Mean	CEEB English Composition Stan. Rating S. D.	Z-Score[1] Mean	Z-Score[1] S. D.	Theme Rating 1 Mean	Theme Rating 1 S. D.	September Theme Theme Rating 2 Mean	September Theme Theme Rating 2 S. D.	Theme Rating Total Mean	Theme Rating Total S. D.
Exp. Pool plus Control Pool	901 to 943	456.27	94.49	99.50	18.51	4.61	1.57	4.41	1.51	9.03	2.52
Exp. Pool	284 to 296	451.70	90.05	98.63	17.75	4.59	1.60	4.46	1.53	9.05	2.57
Control Pool	617 to 647	458.37	96.38	99.90	18.83	4.61	1.55	4.39	1.50	9.02	2.49
Matched Exp. September 1964	224	453.39	82.21	98.96	15.96	4.48	1.42	4.32	1.41	8.80	2.16
Matched Controls September 1964	224	457.36	85.66	99.00	15.91	4.46	1.39	4.33	1.35	8.80	2.16
Exp. Group January 1965	200	456.99	82.00	99.66	15.85	4.53	1.35	4.36	1.43	8.88	2.13
Control Group January 1965	200	460.90	85.60	99.72	15.82	4.49	1.37	4.38	1.37	8.88	2.13
Exp. Group May 1965	137	463.94	81.64	101.17	15.81	4.58	1.28	4.34	1.43	8.92	2.03
Control Group May 1965	137	468.53	84.36	101.27	15.78	4.48	1.37	4.44	1.32	8.92	2.03
Exp. Group May 1966	43	473.30	80.17	103.60	14.62	Not available				9.26	1.94
Control Group May 1966	43	480.86	80.05	103.67	14.76	"				9.26	1.94

TABLE D-II(1)

PERFORMANCE OF AVAILABLE SAMPLES OF MATCHED PAIRS OF STUDENTS ON THREE CRITERION MEASURES AT BEGINNING, MIDDLE, AND END OF FIRST YEAR AND END OF SECOND YEAR OF COLLEGE: UNIVERSITY 1

Subgroup	N	Time of Testing	Cooperative English Tests: English Expression (1960) Converted Score Mean	S. D.	CEEB English Composition Test Standard Rating Mean	S. D.	Theme Rating 1 Mean	S. D.	Theme Rating 2 Mean	S. D.	Theme Rating Total Mean	S. D.
Experimental	224	Sept. 1964	160.78	7.30	453.39	82.21	4.48	1.42	4.32	1.41	8.80	2.16
Control	224	"	160.49	7.36	457.36	85.66	4.46	1.39	4.33	1.35	8.80	2.16
Experimental	200	Jan. 1965	163.42	6.78	500.37	76.72	4.74	1.56	4.77	1.39	9.51	2.39
Control	200	"	164.56	7.24	507.05	72.20	5.42	1.50	4.84	1.46	10.26	2.40
Experimental	137	May 1965	165.16	6.89	508.61	70.61	4.73	1.32	4.63	1.29	9.36	2.20
Control	137	"	166.53	6.94	527.71	78.43	4.84	1.39	4.59	1.28	9.43	2.19
Experimental	43	May 1966	167.88	7.52	536.91	79.76	4.65	1.27	4.51	1.62	9.16	2.46
Control	43	"	168.47	8.73	553.56	71.17	4.72	1.51	4.40	1.28	9.12	2.44

TABLE D-III(1)

THE PERFORMANCE OF 200 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 1

Subgroup	N	Cooperative Eng. Test Converted Score				Difference in Means, Jan. minus September	r	t-Ratio	Difference in Sept. Means Control minus Experimental			Difference in Jan. Means Control minus Experimental			Degrees of Freedom
		September Mean	S. D.	January Mean	S. D.				Diff.	r	t	Diff.	r	t	
Experimental	200	161.05	7.29	163.42	6.78	2.37	0.67	5.793*	-0.26	0.80	0.796	1.14	0.58	2.515*	199
Control	200	160.79	7.32	164.56	7.24	3.78	0.66	8.816*							
Exp.Males	65	157.49	6.85	161.43	6.79	3.94	0.65	5.544*	0.02	0.75	0.025	-0.09	0.60	0.118	64
Cont.Males	65	157.51	6.99	161.34	7.22	3.83	0.53	4.435*							
Exp.Females	135	162.76	6.87	164.37	6.57	1.61	0.65	3.302*	-0.39	0.79	1.025	1.74	0.54	3.143*	134
Cont.Females	135	162.36	6.94	166.11	6.73	3.75	0.67	7.785*							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	5.27	5.054*	2.94	2.917*	198
Controls	4.85	4.597*	4.77	4.565*	198

*Significant at 0.05 level (two-tailed test).

TABLE D-IV(1)

THE PERFORMANCE OF 200 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 1

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in September Means, Control minus Experimental			Diff. in January Means, Control minus Experimental			
		September		January											
Subgroup	N	Mean	S. D.	Mean	S. D.	Jan. minus September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	200	456.99	82.00	500.37	76.72	43.38	0.59	8.521*	3.91	0.81	1.078	6.68	0.44	1.193	199
Control	200	460.90	85.60	507.05	72.20	46.15	0.60	9.122*							
Exp.Males	65	425.28	76.80	472.35	70.25	47.08	0.48	4.997*	-0.06	0.78	0.009	7.14	0.26	0.738	64
Cont.Males	65	425.22	86.02	479.49	55.45	54.28	0.62	6.452*							
Exp.Females	135	472.26	80.03	513.86	76.06	41.60	0.60	6.882*	5.81	0.81	1.362	6.46	0.45	0.938	134
Cont.Females	135	478.07	79.90	520.32	75.51	42.24	0.56	6.704*							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	46.98	3.920*	41.51	3.686*	198
Controls	52.85	4.251*	40.83	3.864*	198

*Significant at 0.05 level (two-tailed test).

TABLE D-V(1)

THE PERFORMANCE OF 200 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 1

Subgroup	N	Degrees of Freedom	September Theme Total Mean	September Theme Total S. D.	Difference in September Means Control minus Experimental Diff.	r	t-Ratio	January Theme Total Mean	January Theme Total S. D.	Difference in January Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	200	199	8.88	2.13	0.00	1.00	0.000	9.50	2.39	0.75	0.28	3.684*
Controls	200		8.88	2.13				10.26	2.40			
Exp.Males	65	64	7.92	2.28	0.00	1.00	0.000	8.18	2.40	1.09	0.25	2.897*
Cont.Males	65		7.92	2.28				9.28	2.51			
Exp.Females	135	134	9.34	1.88	0.00	1.00	0.000	10.14	2.10	0.59	0.16	2.433*
Cont.Females	135		9.34	1.88				10.73	2.19			

	September Diff. in Means Female minus Male Diff.	t-Ratio	January Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	1.42	4.623*	1.96	5.854*	198
Controls	1.42	4.623*	1.45	4.156*	198

*Significant at 0.05 level (two-tailed test).

TABLE D-VI(1)

THE PERFORMANCE OF 137 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 1

Subgroup	N	Cooperative Eng. Test Converted Score September Mean	September S. D.	January Mean	January S. D.	Difference in Means, Jan. minus September	r	t-Ratio	Difference in Sept. Means Control minus Experimental Diff.	r	t	Difference in Jan. Means Control minus Experimental Diff.	r	t	Degrees of Freedom
Experimental	137	161.70	7.26	163.69	6.70	1.99	0.72	4.404*	-0.30	0.80	0.746	1.99	0.65	4.062*	136
Control	137	161.40	7.53	165.68	6.92	4.28	0.69	8.763*							
Exp.Males	34	157.65	6.53	161.09	6.92	3.44	0.68	3.674*	-0.15	0.74	0.166	1.15	0.68	1.168	33
Cont.Males	34	157.50	7.41	162.24	7.17	4.74	0.46	3.586*							
Exp.Females	103	163.04	6.98	164.55	6.39	1.51	0.71	2.968*	-0.35	0.79	0.780	2.26	0.61	4.013*	102
Cont.Females	103	162.69	7.11	166.82	6.44	4.13	0.74	8.504*							

	September Diff. in Means Female minus Male Diff.	t-Ratio	January Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	5.39	3.937*	3.46	2.665*	135
Controls	5.19	3.625*	4.58	3.467*	135

*Significant at 0.05 level (two-tailed test).

TABLE D-VII(1)

THE PERFORMANCE OF 137 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 1

Subgroup	N	Cooperative Eng. Test Converted Score January Mean	S. D.	May Mean	S. D.	Difference in Means, May minus January	r	t-Ratio	Difference in Jan. Means Control minus Experimental Diff.	r	t	Difference in May Means Control minus Experimental Diff.	r	t	Degrees of Freedom
Experimental	137	163.69	6.70	165.16	6.89	1.47	0.67	3.085*	1.99	0.65	4.062*	1.36	0.48	2.258*	136
Control	137	165.68	6.92	166.53	6.94	0.85	0.72	1.910							
Exp.Males	34	161.09	6.92	160.59	6.73	-0.50	0.53	0.432	1.15	0.68	1.168	1.82	0.19	1.174	33
Cont.Males	34	162.24	7.17	162.41	7.23	0.18	0.73	0.191							
Exp.Females	103	164.55	6.39	166.67	6.25	2.12	0.69	4.309*	2.26	0.61	4.013*	1.21	0.49	1.945	102
Cont.Females	103	166.82	6.44	167.88	6.27	1.07	0.68	2.114*							

	January Diff. in Means Female minus Male Diff.	t-Ratio	May Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	3.46	2.665*	6.08	4.790*	135
Controls	4.58	3.467*	5.47	4.211*	135

*Significant at 0.05 level (two-tailed test).

TABLE D-VIII(1)

THE PERFORMANCE OF 137 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 1

		Cooperative Eng. Test Converted Score				Difference in Means, May minus September			Difference in Sept. Means Control minus Experimental			Difference in May Means Control minus Experimental			Degrees of Freedom
		September		May											
Subgroup	N	Mean	S. D.	Mean	S. D.		r	t-Ratio	Diff.	r	t	Diff.	r	t	
Experimental	137	161.70	7.26	165.16	6.89	3.46	0.71	7.477*	-0.30	0.80	0.746	1.37	0.48	2.258*	136
Control	137	161.40	7.53	166.53	6.94	5.12	0.71	10.859*							
Exp.Males	34	157.65	6.53	160.59	6.73	2.94	0.61	2.875*	-0.15	0.74	0.166	1.82	0.19	1.174	33
Cont.Males	34	157.50	7.41	162.41	7.23	4.91	0.57	4.178*							
Exp.Females	103	163.04	6.98	166.67	6.25	3.63	0.69	7.029*	-0.35	0.79	0.780	1.21	0.49	1.945	102
Cont.Females	103	162.69	7.11	167.88	6.27	5.19	0.73	10.449*							

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	5.39	3.937*	6.08	4.790*	135
Controls	5.19	3.625*	5.47	4.211*	135

*Significant at 0.05 level (two-tailed test).

TABLE D-IX(1)

THE PERFORMANCE OF 137 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 1

Subgroup	N	College Entrance Examination Board Standard Rating September Mean	S. D.	January Mean	S. D.	Difference in Means Jan. minus September	r	t-Ratio	Diff. in September Means, Control minus Experimental Diff.	r	t-Ratio	Diff. in January Means, Control minus Experimental Diff.	r	t-Ratio	d.f.
Experimental	137	463.94	81.64	510.23	74.53	46.29	0.61	7.768*	4.58	0.81	1.054	0.58	0.40	0.083	136
Control	137	468.53	84.36	510.82	74.03	42.29	0.56	6.576*							
Exp.Males	34	429.88	74.85	486.41	62.27	56.53	0.65	5.530*	2.18	0.77	0.227	-7.18	0.20	0.565	33
Cont.Males	34	432.06	85.62	479.24	52.62	47.18	0.63	4.067*							
Exp.Females	103	475.18	80.67	518.10	76.54	42.91	0.58	5.987*	5.38	0.81	1.106	3.15	0.40	0.377	102
Cont.Females	103	480.56	80.38	521.24	77.05	40.68	0.52	5.301*							

	September Diff. in Means Female minus Male Diff.	t-Ratio	January Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	45.30	2.868*	31.69	2.171*	135
Controls	48.50	2.979*	42.00	2.938*	135

*Significant at 0.05 level (two-tailed test).

TABLE D-X(1)

THE PERFORMANCE OF 137 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 1

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in January Means, Control minus Experimental			Diff. in May Means, Control minus Experimental			
		January		May		May minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	January	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	137	510.23	74.53	508.61	70.61	-1.63	0.58	0.285	0.58	0.40	0.083	19.10	0.39	2.695*	136
Control	137	510.82	74.03	527.71	78.43	16.89	0.62	2.960*							
Exp.Males	34	486.41	62.27	478.24	45.78	-8.18	0.43	0.790	-7.18	0.20	0.565	4.65	0.11	0.282	33
Cont.Males	34	479.24	52.62	482.88	88.22	3.65	0.65	0.312							
Exp.Females	103	518.10	76.54	518.63	74.40	0.53	0.59	0.078	3.15	0.40	0.377	23.87	0.41	3.102*	102
Cont.Females	103	521.24	77.05	542.50	68.78	21.26	0.60	3.263*							

	January Diff. in Means Female minus Male		May Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	31.69	2.171*	40.39	2.963*	135
Controls	42.00	2.938*	59.62	4.039*	135

*Significant at 0.05 level (two-tailed test).

TABLE D-XI(1)

THE PERFORMANCE OF 137 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 1

		College Entrance Examination Board Standard Rating							Diff. in September Means, Control minus Experimental			Diff. in May Means, Control minus Experimental			
Subgroup	N	September Mean	S. D.	May Mean	S. D.	Difference in Means May minus September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	137	463.94	81.64	508.61	70.61	44.66	0.59	7.519*	4.58	0.81	1.054	19.10	0.39	2.695*	136
Control	137	468.53	84.36	527.71	78.43	59.18	0.66	10.323*							
Exp.Males	34	429.88	74.85	478.24	45.78	48.35	0.46	4.129*	2.18	0.77	0.227	4.65	0.11	0.282	33
Cont.Males	34	432.06	85.62	482.88	88.22	50.82	0.56	3.591*							
Exp.Females	103	475.18	80.67	518.63	74.40	43.45	0.60	6.280*	5.38	0.81	1.106	23.87	0.41	3.102*	102
Cont.Females	103	480.56	80.38	542.50	68.78	61.94	0.67	10.238*							

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	45.30	2.863*	40.39	2.963*	135
Controls	48.50	2.979*	59.62	4.039*	135

*Significant at 0.05 level (two-tailed test).

TABLE D-XII(1)

THE PERFORMANCE OF 137 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 1

Subgroup	N	Degrees of Freedom	September Theme Total Mean	September Theme Total S. D.	Difference in September Means Control minus Experimental Diff.	r	t-Ratio	January Theme Total Mean	January Theme Total S. D.	Difference in January Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	137	136	8.92	2.03	0.00	1.00	0.000	9.78	2.39	0.81	0.24	3.369*
Controls	137		8.92	2.03				10.60	2.18			
Exp.Males	34	33	7.97	2.08	0.00	1.00	0.000	8.18	2.47	1.76	0.31	3.637*
Cont.Males	34		7.97	2.08				9.94	2.26			
Exp.Females	103	102	9.23	1.91	0.00	1.00	0.000	10.31	2.12	0.50	0.14	1.839
Cont.Females	103		9.23	1.91				10.82	2.11			

4

	September Diff. in Means Female minus Male Diff.	September t-Ratio	January Diff. in Means Female minus Male Diff.	January t-Ratio	Degrees of Freedom
Experimentals	1.26	3.248*	2.18	4.847*	135
Controls	1.26	3.248*	0.88	2.044*	135

*Significant at 0.05 level (two-tailed test).

TABLE D-XIII(1)

THE PERFORMANCE OF 137 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 1

Subgroup	N	Degrees of Freedom	January Theme Mean	January Total S. D.	Difference in January Means Control minus Experimental Diff.	r	t-Ratio	May Theme Mean	May Total S. D.	Difference in May Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	137	136	9.78	2.39	0.81	0.24	3.369*	9.36	2.20	0.07	0.16	0.300
Controls	137		10.60	2.18				9.43	2.19			
Exp.Males	34	33	8.18	2.47	1.76	0.31	3.637*	8.41	2.03	0.50	0.15	1.042
Cont.Males	34		9.94	2.26				8.91	2.20			
Exp.Females	103	102	10.31	2.12	0.50	0.14	1.839	9.67	2.16	-0.07	0.13	0.241
Cont.Females	103		10.82	2.11				9.60	2.16			

	January Diff. in Means Female minus Male Diff.	t-Ratio	May Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	2.13	4.847*	1.26	2.961*	135
Controls	0.88	2.044*	0.69	1.598	135

*Significant at 0.05 level (two-tailed test).

TABLE D-XIV(1)

THE PERFORMANCE OF 137 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 1

Subgroup	N	Degrees of Freedom	September Theme Mean	September Total S. D.	Difference in September Means Control minus Experimental Diff.	r	t-Ratio	May Theme Mean	May Total S. D.	Difference in May Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	137	136	8.92	2.03	0.00	1.00	0.000	9.36	2.20	0.07	0.16	0.300
Controls	137		8.92	2.03				9.43	2.19			
Exp.Males	34	33	7.97	2.08	0.00	1.00	0.000	8.41	2.03	0.50	0.15	1.042
Cont.Males	34		7.97	2.08				8.91	2.20			
Exp.Females	103	102	9.23	1.91	0.00	1.00	0.000	9.67	2.16	-0.07	0.13	0.241
Cont.Females	103		9.23	1.91				9.60	2.16			

	September Diff. in Means Female minus Male Diff.	t-Ratio	May Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	1.26	3.248*	1.26	2.961*	135
Controls	1.26	3.248*	0.69	1.598	135

*Significant at 0.05 level (two-tailed test).

TABLE D-XV(1)

THE PERFORMANCE OF 43 MATCHED PAIRS OF STUDENTS ON COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964, JANUARY 1965, MAY 1965, AND MAY 1966; INCLUDIN DIFFERENCES IN MEANS, STANDARD DEVIATIONS, AND t-RATIOS: UNIVERSITY 1

Sub-group	Cooperative English Test Converted Score Sept. 1964 Mean	S. D.	Jan. 1965 Mean	S. D.	Diff. in Means, Jan. minus Sept.	r	t-Ratio df=42	Diff. in Jan. Means, Cont. minus Exper.	r	t-Ratio df=42
Exper.	162.91	6.76	164.56	6.05	1.65	0.74	2.274*	2.40	0.65	2.795*
Control	162.30	6.78	166.95	7.09	4.65	0.71	5.730*			

Sub-group	Cooperative English Test Converted Score Jan. 1965 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Jan.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	164.56	6.05	166.63	6.30	2.07	0.64	2.555*	1.07	0.41	0.945
Control	166.95	7.09	167.70	7.10	0.74	0.72	0.908			

Sub-group	Cooperative English Test Converted Score Sept. 1964 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	162.91	6.76	166.63	6.30	3.72	0.70	4.771*	Same as above		
Control	162.30	6.78	167.70	7.10	5.40	0.75	7.048*			

Sub-group	Cooperative English Test Converted Score May 1965 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus May	r	t-Ratio	Diff. in May 1966 Means Cont. minus Exper.	r	t-Ratio
Exper.	166.63	6.30	167.88	7.52	1.26	-0.06	0.805	0.58	-0.13	0.308
Control	167.70	7.10	168.47	8.73	0.77	0.07	0.458			

Sub-group	Cooperative English Test Converted Score Sept. 1964 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May 1966 Means Cont. minus Exper.	r	t-Ratio
Exper.	162.91	6.76	167.88	7.52	4.98	0.03	3.236*	Same as above		
Control	162.30	6.78	168.47	8.73	6.16	0.29	4.250*			

*Significant at 0.05 level (two-tailed test).

TABLE D-XVI(1)

THE PERFORMANCE OF 43 MATCHED PAIRS OF STUDENTS ON THE COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TEST IN SEPTEMBER 1964, JANUARY 1965, MAY 1965, AND MAY 1966 INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, AND t-RATIOS: UNIVERSITY 1

Sub-group	College Entrance Exam. Board Stan. Rating				Diff. in Means, Jan. minus Sept.	r	t-Ratio df=42	Diff. in Jan. Means, Cont. minus Exper.	r	t-Ratio df=42
	Sept. 1964 Mean	Sept. 1964 S. D.	Jan. 1965 Mean	Jan. 1965 S. D.						
Exper.	473.30	80.17	524.05	74.83	50.74	0.66	5.108*	- 5.21	0.40	0.384
Control	480.86	80.05	518.84	84.42	37.98	0.51	3.013*			

Sub-group	College Entrance Exam. Board Stan. Rating				Diff. in Means, May minus Jan.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
	Jan. 1965 Mean	Jan. 1965 S. D.	May 1965 Mean	May 1965 S. D.						
Exper.	524.05	74.83	523.63	67.17	- 0.42	0.59	0.042	26.79	0.36	2.049*
Control	518.84	84.42	550.42	81.01	31.58	0.66	2.990*			

Sub-group	College Entrance Exam. Board Stan. Rating				Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
	Sept. 1964 Mean	Sept. 1964 S. D.	May 1965 Mean	May 1965 S. D.						
Exper.	473.30	80.17	523.63	67.17	50.33	0.56	4.655*	Same as above		
Control	480.86	80.05	550.42	81.01	69.56	0.64	6.581*			

Sub-group	College Entrance Exam. Board Stan. Rating				Diff. in Means, May minus May	r	t-Ratio	Diff. in May 1966 Means, Cont. minus Exper.	r	t-Ratio
	May 1965 Mean	May 1965 S. D.	May 1966 Mean	May 1966 S. D.						
Exper.	523.63	67.17	536.91	79.76	13.28	0.61	1.308	16.65	0.54	1.482
Control	550.42	81.01	553.56	71.17	3.14	0.70	0.340			

Sub-group	College Entrance Exam. Board Stan. Rating				Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May 1966 Means, Cont. minus Exper.	r	t-Ratio
	Sept. 1964 Mean	Sept. 1964 S. D.	May 1966 Mean	May 1966 S. D.						
Exper.	473.30	80.17	536.91	79.76	63.60	0.59	5.717*	Same as above		
Control	480.86	80.05	553.56	71.17	72.70	0.59	6.803*			

*Significant at 0.05 level (two-tailed test).

TABLE D-XVII(1)

THE PERFORMANCE OF 43 MATCHED PAIRS OF STUDENTS ON THE TOTAL OF TWO THEME RATINGS IN SEPTEMBER 1964, JANUARY 1965, MAY 1965, AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, AND t-RATIOS: UNIVERSITY 1

		Theme Rating 1		Theme Rating 2		Theme Rating Total						
Date	Subgroup	Mean	S. D.	Mean	S. D.	Mean	S. D.	r, Reader 1 with Reader 2	Diff. in Means on Theme Total Cont.minus Exp.	r, Cont. with Exp.	t-Ratio	d.f.
9/64	Exper.	Not available				9.26	1.94	Not available	0.00	1.00	0.000	42
9/64	Control					9.26	1.94					
1/65	Exper.	5.02	1.39	5.19	1.47	10.21	2.33	0.33	0.42	0.31	1.074	42
1/65	Cont.	5.60	1.31	5.02	1.30	10.63	1.93	0.09				
5/65	Exper.	4.93	1.45	4.60	1.22	9.53	2.26	0.42	0.09	0.05	0.207	42
5/65	Cont.	4.88	1.47	4.74	1.14	9.63	1.97	0.12				
5/66	Exper.	4.65	1.27	4.51	1.62	9.16	2.46	0.44	-0.05	0.12	0.093	42
5/66	Cont.	4.72	1.51	4.40	1.28	9.11	2.44	0.53				

LIST OF TABLES
University 2

Table		Page
D-I(2)	Achievement as of September 1964 of a Sample of 1964-65 Freshman Students and of Various Persisting Portions of that Sample: University 2	217
D-II(2)	Performance of Available Samples of Matched Pairs of Students on Three Criterion Measures at Beginning, Middle, and End of First Year and End of Second Year of College: University 2	219
D-III(2)	The Performance of 127 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 2	220
D-IV(2)	The Performance of 127 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 2	221
D-V(2)	The Performance of 127 Matched Pairs of Students on Total Theme Rating in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 2	222
D-VI(2)	The Performance of 98 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 2	223
D-VII(2)	The Performance of 98 Matched Pairs of Students on the Cooperative English Tests: English Expression in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 2	224

Table Page

D-VIII(2) The Performance of 98 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 2 225

D-IX(2) The Performance of 98 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 2 226

D-X(2) The Performance of 98 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 2 227

D-XI(2) The Performance of 98 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 2 228

D-XII(2) The Performance of 98 Matched Pairs of Students on Total Theme Rating in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 2 229

D-XIII(2) The Performance of 98 Matched Pairs of Students on Total Theme Rating in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 2 230

D-XIV(2) The Performance of 98 Matched Pairs of Students on Total Theme Rating in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 2 231

Table Page

D-XV(2) The Performance of 23 Matched Pairs of Students on Cooperative English Tests: English Expression in September 1964, January 1965, May 1965, and May 1966; Including Differences in Means, Standard Deviations, and t-Ratios: University 2 232

D-XVI(2) The Performance of 23 Matched Pairs of Students on the College Entrance Examination Board English Composition Test in September 1964, January 1965, May 1965, and May 1966; Including Differences in Means, Standard Deviations, and t-Ratios: University 2 233

D-XVII(2) The Performance of 23 Matched Pairs of Students on the Total of Two Theme Ratings in September 1964, January 1965, May 1965, and May 1966; Including Differences in Means, Standard Deviations, and t-Ratios: University 2 234

TABLE D-I(2)

ACHIEVEMENT AS OF SEPTEMBER 1964 OF A SAMPLE OF 1964-65 FRESHMAN STUDENTS AND OF VARIOUS PERSISTING PORTIONS OF THAT SAMPLE: UNIVERSITY 2

Sample and Subgroups	N	Percent Men	Percentile Rank in H. S. Class Mean	Percentile Rank in H. S. Class S. D.	SAT-Verbal Score Mean	SAT-Verbal Score S. D.	SAT-V+M Score Mean	SAT-V+M Score S. D.	COOP English Exp. Converted Score Mean	COOP English Exp. Converted Score S. D.
Exp. Pool plus Control Pool	821	47.5	70.56	17.16	516.18	91.66	1,048.36	165.98	164.01	8.21
Exp. Pool	280	47.9	69.89	17.10	510.86	91.82	1,035.90	165.57	164.31	8.69
Control Pool	541	47.3	70.91	17.18	518.93	91.45	1,054.80	165.82	163.85	7.95
Matched Exp. September 1964	198	43.4	70.15	16.22	507.07	84.61	1,025.99	153.58	163.73	7.52
Matched Controls September 1964	198	43.4	71.87	15.57	509.01	82.88	1,040.77	149.31	163.95	7.10
Exp. Group January 1965	127	51.1	69.53	16.21	509.13	87.63	1,024.28	166.84	163.57	7.75
Control Group January 1965	127	51.1	71.50	16.46	506.43	86.41	1,046.41	154.22	163.62	7.47
Exp. Group May 1965	98	42.9	72.76	15.17	510.28	86.20	1,027.81	155.57	164.53	6.76
Control Group May 1965	98	42.9	73.57	15.07	519.51	81.18	1,059.02	150.00	164.83	6.67
Exp. Group May 1966	23	56.5	76.96	14.28	Not Available				166.35	6.60
Control Group May 1966	23	56.5	76.52	14.63	"				166.09	6.74

TABLE D-I(2)
CONTINUED

Sample and Subgroups	N	CEEB English Composition Stan. Rating Mean	CEEB English Composition Stan. Rating S. D.	Z-Score[1] Mean	Z-Score[1] S. D.	September Theme: Theme Rating 1 Mean	Theme Rating 1 S. D.	Theme Rating 2 Mean	Theme Rating 2 S. D.	Theme Rating Total Mean	Theme Rating Total S. D.
Exp. Pool plus Control Pool	821	504.32	90.97	99.44	18.39	4.92	1.52	4.66	1.55	9.58	2.58
Exp. Pool	280	504.52	92.72	99.83	19.05	5.04	1.51	4.61	1.66	9.64	2.68
Control Pool	541	504.21	90.06	99.24	18.03	4.86	1.52	4.69	1,49	9.55	2.52
Matched Exp. September 1964	198	500.75	83.73	98.70	16.47	5.06	1.29	4.56	1.45	9.61	2.12
Matched Controls September 1964	198	497.74	85.35	98.65	16.46	5.01	1.36	4.59	1.32	9.61	2.12
Exp. Group January 1965	127	496.92	87.89	98.11	17.35	5.06	1.28	4.65	1.47	9.71	2.17
Control Group January 1965	127	496.16	87.02	98.07	17.30	5.14	1.40	4.55	1.33	9.71	2.17
Exp. Group May 1965	98	516.12	79.38	101.34	15.05	5.30	1.16	4.69	1.41	9.96	1.93
Control Group May 1965	98	513.06	76.41	101.36	14.78	5.32	1.17	4.64	1.30	9.96	1.93
Exp. Group May 1966	23	547.04	74.24	106.91	15.38	Not Available				9.87	2.05
Control Group May 1966	23	548.48	75.01	106.74	14.77	"				9.87	2.05

---

[1] Combination of Cooperative English Tests: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

TABLE D-II(2)

PERFORMANCE OF AVAILABLE SAMPLES OF MATCHED PAIRS OF STUDENTS ON THREE CRITERION MEASURES AT BEGINNING, MIDDLE, AND END OF FIRST YEAR AND END OF SECOND YEAR OF COLLEGE: UNIVERSITY 2

			Cooperative English Tests: English Expression (1960) Converted Score		CEEB English Composition Test Standard Rating		Theme Rating 1		Theme Rating 2		Theme Rating Total	
Subgroup	N	Time of Testing	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.
Experimental	198	Sept. 1964	163.73	7.52	500.75	83.73	5.06	1.29	4.56	1.45	9.61	2.12
Control	198	"	163.95	7.10	497.74	85.35	5.01	1.36	4.59	1.32	9.61	2.12
Experimental	127	Jan. 1965	165.67	7.79	508.49	86.78	5.13	1.47	5.19	1.63	10.31	2.62
Control	127	"	166.50	7.77	511.73	81.43	4.88	1.37	5.16	1.52	10.04	2.44
Experimental	98	May 1965	167.85	6.62	528.17	92.05	5.10	1.27	5.24	1.33	10.35	2.20
Control	98	"	168.27	7.22	537.06	81.52	4.91	1.49	5.48	1.42	10.39	2.21
Experimental	23	May 1966	171.00	8.11	572.22	75.43	5.61	1.31	5.91	1.53	11.52	2.62
Control	23	"	171.30	5.55	558.61	75.24	5.04	1.63	4.87	1.65	9.91	2.75

TABLE D-III(2)

THE PERFORMANCE OF 127 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 2

		Cooperative Eng. Test Converted Score				Difference in Means, Jan. minus September			Difference in Sept. Means Control minus Experimental			Difference in Jan. Means Control minus Experimental			Degrees of Freedom
		September		January											
Subgroup	N	Mean	S. D.	Mean	S. D.		r	t-Ratio	Diff.	r	t	Diff.	r	t	
Experimental	127	163.57	7.75	165.67	7.79	2.10	0.79	4.645*	0.06	0.86	0.152	0.83	0.66	1.444	126
Control	127	163.62	7.47	166.50	7.77	2.87	0.76	6.165*							
Exp.Males	65	162.52	7.63	164.89	7.67	2.37	0.76	3.615*	-0.06	0.86	0.124	0.97	0.67	1.228	64
Cont.Males	65	162.46	7.10	165.86	7.81	3.40	0.72	4.830*							
Exp.Females	62	164.66	7.72	166.48	7.83	1.82	0.80	2.912*	0.18	0.85	0.334	0.68	0.64	0.808	61
Cont.Females	62	164.84	7.65	167.16	7.67	2.32	0.81	3.846*							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	2.14	1.558	1.59	1.148	125
Controls	2.38	1.802	1.30	0.938	125

*Significant at 0.05 level (two-tailed test).

TABLE D-IV(2)

THE PERFORMANCE OF 127 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 2

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in September Means, Control minus Experimental			Diff. in January Means, Control minus Experimental			
		September		January		Jan. minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	127	496.92	87.89	508.49	86.78	11.57	0.76	2.133*	-0.80	0.87	0.200	3.24	0.51	0.438	126
Control	127	496.13	87.02	511.73	81.43	15.61	0.64	2.453*							
Exp.Males	65	487.92	84.13	495.75	84.21	7.83	0.78	1.111	-0.86	0.86	0.148	14.22	0.45	1.341	64
Cont.Males	65	487.06	89.44	509.97	76.25	22.91	0.56	2.333*							
Exp.Females	62	506.35	90.71	521.84	87.42	15.48	0.73	1.860	-0.73	0.88	0.133	-8.26	0.58	0.810	61
Cont.Females	62	505.63	83.36	513.58	86.50	7.95	0.73	1.000							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	18.43	1.179	26.09	1.699	125
Controls	18.57	1.199	3.61	0.248	125

*Significant at 0.05 level (two-tailed test).

TABLE D-V(2)

THE PERFORMANCE OF 127 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 2

Subgroup	N	Degrees of Freedom	September Theme Total Mean	September Theme Total S. D.	Difference in September Means Control minus Experimental Diff.	r	t-Ratio	January Theme Total Mean	January Theme Total S. D.	Difference in January Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	127	126	9.71	2.17	0.00	1.00	0.000	10.31	2.62	-0.28	0.26	1.001
Controls	127		9.71	2.17				10.04	2.44			
Exp.Males	65	64	9.40	2.15	0.00	1.00	0.000	9.75	2.68	-0.18	0.30	0.486
Cont.Males	65		9.40	2.15				9.57	2.44			
Exp.Females	62	61	10.03	2.15	0.00	1.00	0.000	10.90	2.42	-0.37	0.13	0.922
Cont.Females	62		10.03	2.15				10.53	2.35			

	September Diff. in Means Female minus Male Diff.	t-Ratio	January Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.63	1.643	1.15	2.513*	125
Controls	0.63	1.643	0.96	2.246*	125

*Significant at 0.05 level (two-tailed test).

TABLE D-VI(2)

THE PERFORMANCE OF 98 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 2

		Cooperative Eng. Test Converted Score				Difference in Means, Jan. minus			Difference in Sept. Means Control minus Experimental			Difference in Jan. Means Control minus Experimental			
		September		January											
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t	Diff.	r	t	Degrees of Freedom
Experimental	98	164.53	6.76	167.42	7.28	2.89	0.75	5.707*	0.30	0.83	0.735	0.71	0.65	1.219	97
Control	98	164.83	6.67	168.13	6.44	3.31	0.77	7.283*							
Exp.Males	42	163.33	7.43	166.38	7.50	3.05	0.78	3.941*	-0.12	0.86	0.203	0.36	0.64	0.380	41
Cont.Males	42	163.21	6.76	166.74	6.46	3.52	0.69	4.307*							
Exp.Females	56	165.43	6.05	168.20	7.01	2.77	0.72	4.102*	0.61	0.78	1.101	0.98	0.65	1.309	55
Cont.Females	56	166.04	6.34	169.18	6.22	3.14	0.82	6.156*							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	2.10	1.521	1.82	1.218	96
Controls	2.83	2.097*	2.44	1.871	96

*Significant at 0.05 level (two-tailed test).

TABLE D-VII(2)

THE PERFORMANCE OF 98 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 2

Subgroup	N	Cooperative Eng. Test Converted Score				Difference in Means, May minus January	r	t-Ratio	Difference in Jan. Means Control minus Experimental			Difference in May Means Control minus Experimental			Degrees of Freedom
		January Mean	S. D.	May Mean	S. D.				Diff.	r	t	Diff.	r	t	
Experimental	98	167.42	7.28	167.85	6.62	0.43	0.70	0.780	0.71	0.65	1.219	0.42	0.54	0.617	97
Control	98	168.13	6.44	168.27	7.22	0.13	0.69	0.243							
Exp.Males	42	166.38	7.50	166.95	7.62	0.57	0.73	0.660	0.36	0.64	0.380	-0.29	0.50	0.255	41
Cont.Males	42	166.74	6.46	166.67	6.63	-0.07	0.57	0.075							
Exp.Females	56	168.20	7.01	168.52	5.67	0.32	0.67	0.449	0.98	0.65	1.309	0.95	0.57	1.126	55
Cont.Females	56	169.18	6.22	169.46	7.41	0.29	0.77	0.443							

	January Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	1.82	1.218	1.57	1.155	96
Controls	2.44	1.871	1.79	1.914	96

TABLE D-VIII(2)

THE PERFORMANCE OF 98 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 2

		Cooperative Eng. Test Converted Score September		May		Difference in Means, May minus			Difference in Sept. Means Control minus Experimental			Difference in May Means Control minus Experimental			Degrees of
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t	Diff.	r	t	Freedom
Experimental	98	164.53	6.76	167.85	6.62	3.32	0.79	7.443*	0.30	0.83	0.735	0.42	0.54	0.617	97
Control	98	164.83	6.67	168.27	7.22	3.44	0.71	6.352*							
Exp.Males	42	163.33	7.43	166.95	7.62	3.62	0.79	4.788*	-0.12	0.86	0.203	-0.29	0.50	0.255	41
Cont.Males	42	163.21	6.76	166.67	6.63	3.45	0.62	3.791*							
Exp.Females	56	165.43	6.05	168.52	5.67	3.09	0.77	5.726*	0.61	0.78	1.101	0.95	0.57	1.126	55
Cont.Females	56	166.04	6.34	169.46	7.41	3.43	0.75	5.164*							

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of
	Diff.	t-Ratio	Diff.	t-Ratio	Freedom
Experimentals	2.10	1.521	1.57	1.155	96
Controls	2.83	2.097*	2.79	1.914	96

*Significant at 0.05 level (two-tailed test).

TABLE D-IX(2)

THE PERFORMANCE OF 98 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 2

		College Entrance Examination Board Standard Rating							Diff. in September Means, Control minus Experimental			Diff. in January Means, Control minus Experimental			
		September		January		Difference in Means Jan. minus September									
Subgroup	N	Mean	S. D.	Mean	S. D.		r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	98	516.12	79.38	518.64	81.98	2.52	0.72	0.411	-3.06	0.84	0.675	7.38	0.45	0.860	97
Control	98	513.06	76.41	526.02	79.70	12.96	0.55	1.730							
Exp.Males	42	507.79	78.03	505.24	87.09	-2.55	0.81	0.313	1.42	0.86	0.206	7.40	0.36	0.515	41
Cont.Males	42	509.21	86.31	512.64	75.06	3.43	0.50	0.271							
Exp.Females	56	522.38	79.81	528.70	76.41	6.32	0.65	0.714	-6.43	0.83	1.069	7.35	0.51	0.696	55
Cont.Females	56	515.95	67.90	536.05	81.58	20.10	0.61	2.224*							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	14.59	0.895	23.46	1.402	96
Controls	6.74	0.428	23.41	1.439	96

*Significant at 0.05 level (two-tailed test).

TABLE D-X(2)

THE PERFORMANCE OF 98 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 2

		College Entrance Examination Board Standard Rating				Difference in Means May minus			Diff. in January Means, Control minus Experimental			Diff. in May Means, Control minus Experimental			
		January		May											
Subgroup	N	Mean	S. D.	Mean	S. D.	January	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	98	518.64	81.98	528.17	92.05	9.53	0.70	1.373	7.38	0.45	0.860	8.89	0.46	0.965	97
Control	98	526.02	79.70	537.06	81.52	11.04	0.62	1.540							
Exp.Males	42	505.24	87.09	523.95	90.23	18.71	0.86	2.520*	7.40	0.36	0.515	2.50	0.61	0.214	41
Cont.Males	42	512.64	75.06	526.45	75.94	13.81	0.49	1.156							
Exp.Females	56	528.70	76.41	531.34	93.26	2.64	0.57	0.245	7.35	0.51	0.696	13.68	0.36	1.006	55
Cont.Females	56	536.05	81.58	545.02	84.59	8.97	0.69	1.010							

	January Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	23.46	1.402	7.39	0.389	96
Controls	23.41	1.439	18.57	1.111	96

*Significant at 0.05 level (two-tailed test).

TABLE D-XI(2)

THE PERFORMANCE OF 98 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 2

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in September Means, Control minus Experimental			Diff. in May Means, Control minus Experimental			
		September		May		May minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	98	516.12	79.38	528.17	92.05	12.05	0.70	1.764	-3.06	0.84	0.675	8.89	0.46	0.965	97
Control	98	513.06	76.41	537.06	81.52	24.00	0.55	3.161*							
Exp.Males	42	507.79	78.03	523.95	90.23	16.16	0.79	1.858	1.42	0.86	0.206	2.50	0.61	0.214	41
Cont.Males	42	509.21	86.31	526.45	75.94	17.24	0.51	1.367							
Exp.Females	56	522.38	79.81	531.34	93.26	8.96	0.64	0.891	-6.43	0.83	1.069	13.68	0.36	1.006	55
Cont.Females	56	515.95	67.90	545.02	84.59	29.07	0.60	3.101*							

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	14.59	0.895	7.39	0.389	96
Controls	6.74	0.428	18.57	1.111	96

*Significant at 0.05 level (two-tailed test).

TABLE D-XII(2)

THE PERFORMANCE OF 98 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 2

			September Theme Total		Difference in September Means Control minus Experimental			January Theme Total		Difference in January Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Mean	S. D.	Diff.	r	t-Ratio	Mean	S. D.	Diff.	r	t-Ratio
Experimentals	98	97	9.96	1.93	0.00	1.00	0.000	10.84	2.28	-0.32	0.13	1.041
Controls	98		9.96	1.93				10.52	2.26			
Exp.Males	42	41	9.71	2.00	0.00	1.00	0.000	10.43	2.44	-0.43	0.22	0.945
Cont.Males	42		9.71	2.00				10.00	2.20			
Exp.Females	56	55	10.14	1.85	0.00	1.00	0.000	11.14	2.11	-0.23	0.00	0.564
Cont.Females	56		10.14	1.85				10.91	2.21			

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.43	1.085	0.71	1.535	96
Controls	0.43	1.085	0.91	1.998*	96

*Significant at 0.05 level (two-tailed test).

TABLE D-XIII(2)

THE PERFORMANCE OF 98 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 2

Subgroup	N	Degrees of Freedom	January Theme Mean	January Total S. D.	Difference in January Means Control minus Experimental Diff.	r	t-Ratio	May Theme Mean	May Total S. D.	Difference in May Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	98	97	10.84	2.28	-0.32	0.13	1.041	10.35	2.20	0.04	0.16	0.141
Controls	98		10.52	2.26				10.39	2.21			
Exp.Males	42	41	10.43	2.44	-0.43	0.22	0.945	10.50	2.23	-0.62	0.14	1.301
Cont.Males	42		10.00	2.20				9.88	2.41			
Exp.Females	56	55	11.14	2.11	-0.23	0.00	0.564	10.23	2.16	0.54	0.20	1.525
Cont.Females	56		10.91	2.21				10.77	1.95			

	January Diff. in Means Female minus Male Diff.	January t-Ratio	May Diff. in Means Female minus Male Diff.	May t-Ratio	Degrees of Freedom
Experimentals	0.71	1.535	-0.27	0.593	96
Controls	0.91	1.998*	0.89	1.988*	96

*Significant at 0.05 level (two-tailed test).

TABLE D-XIV(2)

THE PERFORMANCE OF 98 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 2

Subgroup	N	Degrees of Freedom	September Theme Mean	September Total S. D.	Difference in September Means Control minus Experimental Diff.	r	t-Ratio	May Theme Mean	May Total S. D.	Difference in May Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	98	97	9.96	1.93	0.00	1.00	0.000	10.35	2.20	0.04	0.16	0.141
Controls	98		9.96	1.93				10.39	2.21			
Exp.Males	42	41	9.71	2.00	0.00	1.00	0.000	10.50	2.23	-0.62	0.14	1.301
Cont.Males	42		9.71	2.00				9.88	2.41			
Exp.Females	56	55	10.14	1.85	0.00	1.00	0.000	10.23	2.16	0.54	0.20	1.525
Cont.Females	56		10.14	1.85				10.77	1.95			

	September Diff. in Means Female minus Male Diff.	t-Ratio	May Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.43	1.085	-0.27	0.593	96
Controls	0.43	1.085	0.89	1.988*	96

*Significant at 0.05 level (two-tailed test).

TABLE D-XV(2)

THE PERFORMANCE OF 23 MATCHED PAIRS OF STUDENTS ON COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964, JANUARY 1965, MAY 1965, AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, AND t-RATIOS: UNIVERSITY 2

Sub-group	Cooperative English Test Converted Score Sept. 1964 Mean	S. D.	Jan. 1965 Mean	S. D.	Diff. in Means, Jan. minus Sept.	r	t-Ratio df=22	Diff. in Jan. Means, Cont. minus Exper.	r	t-Ratio df=22
Exper.	166.35	6.60	168.74	7.66	2.39	0.80	2.421*	0.87	0.54	0.605
Control	166.09	6.74	169.61	6.20	3.52	0.75	3.558*			

Sub-group	Cooperative English Test Converted Score Jan. 1965 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Jan.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	168.74	7.66	169.48	7.13	0.74	0.67	0.579	1.35	0.65	1.147
Control	169.61	6.20	170.83	5.71	1.22	0.84	1.667			

Sub-group	Cooperative English Test Converted Score Sept. 1964 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	166.35	6.60	169.48	7.13	3.13	0.90	4.657*	Same as above		
Control	166.09	6.74	170.83	5.71	4.74	0.76	5.057*			

Sub-group	Cooperative English Test Converted Score May 1965 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus May	r	t-Ratio	Diff. in May 1966 Means Cont. minus Exper.	r	t-Ratio
Exper.	169.48	7.13	171.00	8.11	1.52	0.88	1.836	0.30	0.51	0.200
Control	170.83	5.71	171.30	5.55	0.48	0.88	0.812			

Sub-group	Cooperative English Test Converted Score Sept. 1964 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May 1966 Means Cont. minus Exper.	r	t-Ratio
Exper.	166.35	6.60	171.00	8.11	4.65	0.87	5.374*	Same as above		
Control	166.09	6.74	171.30	5.55	5.22	0.69	4.960*			

*Significant at 0.05 level (two-tailed test).

TABLE D-XVI(2)

THE PERFORMANCE OF 23 MATCHED PAIRS OF STUDENTS ON THE COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TEST IN SEPTEMBER 1964, JANUARY 1965, MAY 1965, AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, AND t-RATIOS: UNIVERSITY 2

Sub-group	College Entrance Exam. Board Stan. Rating Sept. 1964 Mean	S. D.	Jan. 1965 Mean	S. D.	Diff. in Means, Jan. minus Sept.	r	t-Ratio df=22	Diff. in Jan. Means, Cont. minus Exper.	r	t-Ratio df=22
Exper.	547.04	74.24	542.09	81.08	- 4.96	0.84	0.525	-15.22	0.44	0.785
Control	548.48	75.01	526.87	89.73	-21.61	0.49	1.207			

Sub-group	College Entrance Exam. Board Stan. Rating Jan. 1965 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Jan.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	542.09	81.08	545.69	94.17	3.61	0.76	0.274	6.61	0.37	0.319
Control	526.87	89.73	552.30	77.89	25.43	0.62	1.611			

Sub-group	College Entrance Exam. Board Stan. Rating Sept. 1964 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	547.04	74.24	545.69	94.17	- 1.35	0.72	0.097	Same as above		
Control	548.48	75.01	552.30	77.89	3.83	0.56	0.250			

Sub-group	College Entrance Exam. Board Stan. Rating May 1965 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus May	r	t-Ratio	Diff. in May 1966 Means, Cont. minus Exper.	r	t-Ratio
Exper.	545.69	94.17	572.22	75.43	26.52	0.74	1.968	-13.61	0.59	0.940
Control	552.30	77.89	558.61	75.24	6.30	0.64	0.453			

Sub-group	College Entrance Exam. Board Stan. Rating Sept. 1964 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May 1966 Means, Cont. minus Exper.	r	t-Ratio
Exper.	547.04	74.24	572.22	75.43	25.17	0.74	2.196*	Same as above		
Control	548.48	75.01	558.61	75.24	10.13	0.67	0.784			

*Significant at 0.05 level (two-tailed test).

TABLE D-XVII(2)

THE PERFORMANCE OF 23 MATCHED PAIRS OF STUDENTS ON THE TOTAL OF TWO THEME RATINGS IN SEPTEMBER 1964, JANUARY 1965, MAY 1965, AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, AND t-RATIOS: UNIVERSITY 2

Date	Subgroup	Theme Rating 1		Theme Rating 2		Theme Rating Total		r, Reader 1 with Reader 2	Diff. in Means on Theme Total Cont.minus Exp.	r, Cont. with Exp.	t-Ratio	d.f.
		Mean	S. D.	Mean	S. D.	Mean	S. D.					
9/64	Exper.	Not available				9.87	2.05	Not available	0.00	1.00	0.000	22
9/64	Control					9.87	2.05					
1/65	Exper.	5.78	1.18	5.83	1.27	11.61	1.88	0.18	-0.83	0.29	1.609	22
1/65	Control	5.43	1.10	5.35	1.55	10.78	2.15	0.29				
5/65	Exper.	5.43	1.25	5.52	1.38	10.96	2.27	0.50	-1.09	-0.06	1.483	22
5/65	Control	5.00	1.44	4.87	1.42	9.87	2.44	0.44				
5/66	Exper.	5.61	1.31	5.91	1.53	11.52	2.62	0.70	-1.61	0.37	2.500*	22
5/66	Control	5.04	1.63	4.87	1.65	9.91	2.75	0.41				

*Significant at 0.05 level (two-tailed test).

LIST OF TABLES
University 3

Table		Page
D-I(3)	Achievement as of September 1964 of a sample of 1964-65 Freshman Students and of Various Persisting Portions of that Sample: University 3	239
D-II(3)	Performance of Available Samples of Matched Pairs of Students on Three Criterion Measures at Beginning, Middle, and End of First Year and End of Second Year of College: University 3	241
D-III(3)	The Performance of 124 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 3	242
D-IV(3)	The Performance of 124 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 3	243
D-V(3)	The Performance of 124 Matched Pairs of Students on Total Theme Rating in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 3	244
D-VI(3)	The Performance of 40 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 3	245
D-VII(3)	The Performance of 40 Matched Pairs of Students on the Cooperative English Tests: English Expression in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 3	246

Table Page

D-VIII(3) The Performance of 40 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 3 247

D-IX(3) The Performance of 40 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 3 248

D-X(3) The Performance of 40 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 3 249

D-XI(3) The Performance of 40 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 3 250

D-XII(3) The Performance of 40 Matched Pairs of Students on Total Theme Rating in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 3 251

D-XIII(3) The Performance of 40 Matched Pairs of Students on Total Theme Rating in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 3 252

D-XIV(3) The Performance of 40 Matched Pairs of Students on Total Theme Rating in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 3 253

Table Page

D-XV(3) The Performance of 11 Matched Pairs of Students on Cooperative English Tests: English Expression in September 1964, January 1965, May 1965, and May 1966; Including Differences in Means, Standard Deviations, and t-Ratios: University 3 254

D-XVI(3) The Performance of 11 Matched Pairs of Students on The College Entrance Examination Board English Composition Test in September 1964, January 1965, May 1965, and May 1966; Including Differences in Means, Standard Deviations, and t-Ratios: University 3 255

D-XVII(3) The Performance of 11 Matched Pairs of Students on the Total of Two Theme Ratings in September 1964, January 1965, May 1965, and May 1966; Including Differences in Means, Standard Deviations, and t-Ratios: University 3 256

TABLE D-I(3)

ACHIEVEMENT AS OF SEPTEMBER 1964 OF A SAMPLE OF 1964-65 FRESHMAN STUDENTS AND OF VARIOUS PERSISTING PORTIONS OF THAT SAMPLE: UNIVERSITY 3

Sample and Subgroups	N	Percent Men	Percentile Rank in H. S. Class		ACT English Stan. Score		ACT Composite Stan. Score		COOP English Exp. Converted Score	
			Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.
Exp. Pool plus Control Pool	775	32.4	66.25	19.95	20.97	3.38	21.79	3.65	161.30	8.70
Exp. Pool	280	38.1	63.61	20.25	20.76	3.32	21.60	3.52	161.32	8.57
Control Pool	494	29.2	67.71	19.64	21.08	3.41	21.90	3.71	161.27	8.77
Matched Exp. September 1964	204	33.3	63.63	20.18	20.65	3.13	21.50	3.41	161.09	7.36
Matched Controls September 1964	204	33.3	68.14	18.62	20.79	3.18	21.81	3.54	160.82	7.78
Exp. Group January 1965	124	27.3	67.90	20.09	20.97	3.22	21.75	3.58	163.40	7.11
Control Group January 1965	124	27.3	71.94	17.49	20.84	3.33	22.25	3.61	162.75	7.81
Exp. Group May 1965	40	40.0	66.00	21.66	21.10	3.21	21.80	3.46	162.35	7.76
Control Group May 1965	40	40.0	70.75	19.02	21.20	3.26	22.70	3.13	161.73	8.26
Exp. Group May 1966	11	18.2	79.09	15.05	Not Available				167.18	6.04
Control Group May 1966	11	18.2	71.82	18.98	"				166.73	7.11

TABLE D-I(3)
CONTINUED

Sample and Subgroups	N	CEEB English Composition Stan. Rating Mean	S. D.	Z-Score[1] Mean	S. D.	Theme Rating 1 Mean	S. D.	September Theme Theme Rating 2 Mean	S. D.	Theme Rating Total Mean	S. D.
Exp. Pool plus Control Pool	775	467.90	90.22	99.51	18.35	4.33	1.60	4.17	1.47	8.50	2.62
Exp. Pool	280	463.35	92.31	99.03	18.57	4.18	1.60	4.17	1.49	8.35	2.62
Control Pool	494	470.25	88.36	99.74	18.21	4.42	1.59	4.17	1.47	8.59	2.62
Matched Exp. September 1964	204	461.93	83.66	98.61	15.99	4.19	1.48	4.25	1.35	8.44	2.26
Matched Controls September 1964	204	465.08	78.65	98.64	15.79	4.37	1.41	4.07	1.37	8.44	2.26
Exp. Group January 1965	124	483.73	81.17	101.87	14.82	4.67	1.53	4.47	1.38	9.11	2.30
Control Group January 1965	124	489.53	73.76	101.81	14.64	4.62	1.41	4.49	1.45	9.11	2.30
Exp. Group May 1965	40	469.00	78.03	100.80	15.52	4.50	1.34	4.40	1.43	8.90	2.21
Control Group May 1965	40	474.95	69.86	100.75	15.26	4.60	1.53	4.30	1.21	8.90	2.21
Exp. Group May 1966	11	496.91	85.62	109.45	15.55	Not Available				9.09	1.44
Control Group May 1966	11	497.91	82.82	109.09	15.35	"				9.09	1.44

---

[1] Combination of Cooperative English Tests: English Expression and College Entrance Examination Board Engli
Composition Test. September scores. See page 39.

TABLE D-II(3)

PERFORMANCE OF AVAILABLE SAMPLES OF MATCHED PAIRS OF STUDENTS ON THREE CRITERION MEASURES AT BEGINNING, MIDDLE, AND END OF FIRST YEAR AND END OF SECOND YEAR OF COLLEGE: UNIVERSITY 3

Subgroup	N	Time of Testing	Cooperative English Tests: English Expression (1960) Converted Score		CEEB English Composition Test Standard Rating		Theme Rating 1		Theme Rating 2		Theme Rating Total	
			Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.
Experimental	204	Sept.1964	161.09	7.36	461.93	83.66	4.19	1.48	4.25	1.35	8.44	2.26
Control	204	"	160.82	7.78	465.08	78.65	4.37	1.41	4.07	1.37	8.44	2.26
Experimental	124	Jan. 1965	164.59	7.93	502.33	78.10	4.52	1.31	4.59	1.44	9.10	2.32
Control	124	"	165.57	7.84	507.88	83.58	4.47	1.39	4.87	1.64	9.34	2.57
Experimental	40	May 1965	166.15	7.26	497.85	91.03	4.18	1.58	4.23	1.41	8.40	2.47
Control	40	"	166.23	7.76	495.55	79.69	4.33	1.10	4.55	1.34	8.88	2.14
Experimental	11	May 1966	166.82	9.59	567.55	94.07	3.91	1.08	4.64	0.77	8.55	1.44
Control	11	"	170.55	6.85	558.82	79.13	3.82	1.40	3.91	0.90	7.73	2.05

TABLE D-III(3)

THE PERFORMANCE OF 124 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 3

Subgroup	N	Cooperative Eng. Test Converted Score September Mean	September S. D.	January Mean	January S. D.	Difference in Means, Jan. minus September	r	t-Ratio	Difference in Sept. Means Control minus Experimental Diff.	r	t	Difference in Jan. Means Control minus Experimental Diff.	r	t	Degrees of Freedom
Experimental	124	163.40	7.11	164.59	7.93	1.19	0.69	2.203*	-0.65	0.81	1.538	0.98	0.54	1.450	123
Control	124	162.76	7.81	165.57	7.84	2.81	0.75	5.644*							
Exp.Males	34	158.82	7.84	159.56	7.60	0.74	0.53	0.563	-2.09	0.77	2.222*	1.06	0.23	0.713	33
Cont.Males	34	156.74	8.19	160.62	6.00	3.88	0.67	3.679*							
Exp.Females	90	165.13	5.97	166.49	7.19	1.36	0.70	2.431*	-0.10	0.77	0.224	0.96	0.54	1.269	89
Cont.Females	90	165.03	6.31	167.44	7.63	2.41	0.73	4.325*							

	September Diff. in Means Female minus Male Diff.	t-Ratio	January Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	6.31	4.759*	6.93	4.675*	122
Controls	8.29	5.947*	6.82	4.658*	122

*Significant at 0.05 level (two-tailed test).

TABLE D-IV(3)

THE PERFORMANCE OF 124 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 3

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in September Means, Control minus Experimental			Diff. in January Means, Control minus Experimental			
		September		January											
Subgroup	N	Mean	S. D.	Mean	S. D.	Jan. minus September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	124	483.73	81.17	502.33	78.10	18.60	0.64	3.063*	5.80	0.63	1.329	5.55	0.47	0.736	123
Control	124	489.53	73.76	507.88	83.58	18.35	0.63	2.995*							
Exp.Males	34	424.24	67.72	458.03	67.39	33.79	0.64	3.368*	21.06	0.62	2.060*	16.09	0.41	1.397	33
Cont.Males	34	445.29	66.32	474.12	52.72	28.82	0.46	2.642*							
Exp.Females	90	506.21	74.18	519.07	75.27	12.86	0.56	1.737	0.03	0.83	0.007	1.57	0.43	0.166	89
Cont.Females	90	506.24	69.40	520.63	89.34	14.39	0.64	1.955							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	81.97	5.574*	61.04	4.109*	123
Controls	60.95	4.380*	46.51	2.831*	123

*Significant at 0.05 level (two-tailed test).

TABLE D-V(3)

THE PERFORMANCE OF 124 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 3

			September Theme Total		Difference in September Means Control minus Experimental			January Theme Total		Difference in January Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Mean	S. D.	Diff.	r	t-Ratio	Mean	S. D.	Diff.	r	t-Ratio
Experimentals	124	123	9.11	2.30	0.00	1.00	0.000	9.10	2.32	0.23	0.31	0.901
Controls	124		9.11	2.30				9.34	2.57			
Exp.Males	34	33	7.91	2.33	0.00	1.00	0.000	7.50	2.02	0.21	0.07	0.391
Cont.Males	34		7.91	2.33				7.71	2.40			
Exp.Females	90	89	9.57	2.12	0.00	1.00	0.000	9.71	2.14	0.24	0.22	0.817
Cont.Females	90		9.57	2.12				9.96	2.36			

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	1.66	3.742*	2.21	5.178*	122
Controls	1.66	3.742*	2.25	4.676*	122

*Significant at 0.05 level (two-tailed test).

TABLE D-VI(3)

THE PERFORMANCE OF 40 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 3

		Cooperative Eng. Test Converted Score September		January		Difference in Means, Jan. minus			Difference in Sept. Means Control minus Experimental			Difference in Jan. Means Control minus Experimental			Degrees of
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t	Diff.	r	t	Freedom
Experimental	40	162.35	7.76	162.48	8.41	0.13	0.72	0.127	-0.63	0.79	0.742	2.10	0.39	1.413	39
Control	40	161.73	8.26	164.58	8.44	2.85	0.73	2.877*							
Exp.Males	16	159.13	7.65	158.56	7.04	-0.56	0.69	0.375	-2.56	0.68	1.657	2.44	0.36	1.267	15
Cont.Males	16	156.56	7.22	161.00	6.10	4.44	0.59	2.816*							
Exp.Females	24	164.50	7.06	165.08	8.23	0.58	0.67	0.444	0.67	0.82	0.754	1.87	0.28	0.872	23
Cont.Females	24	165.17	7.02	166.96	8.93	1.79	0.74	1.429							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		Degrees of
	Diff.	t-Ratio	Diff.	t-Ratio	Freedom
Experimentals	5.37	2.224*	6.52	2.532*	38
Controls	8.61	3.661*	5.96	2.272*	38

*Significant at 0.05 level (two-tailed test).

TABLE D-VII(3)

THE PERFORMANCE OF 40 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 3

Subgroup	N	Cooperative Eng. Test Converted Score January Mean	S. D.	May Mean	S. D.	Difference in Means, May minus January	r	t-Ratio	Difference in Jan. Means Control minus Experimental Diff.	r	t	Difference in May Means Control minus Experimental Diff.	r	t	Degrees of Freedom
Experimental	40	162.48	8.41	166.15	7.26	3.67	0.74	3.994*	2.10	0.39	1.413	0.07	0.45	0.060	39
Control	40	164.58	8.44	166.23	7.76	1.65	0.75	1.805							
Exp. Males	16	158.56	7.04	162.00	6.64	3.44	0.69	2.478*	2.44	0.36	1.267	0.50	0.00	0.221	15
Cont.Males	16	161.00	6.10	162.50	5.69	1.50	0.56	1.048							
Exp.Females	24	165.08	8.23	168.92	6.26	3.83	0.69	3.076*	1.87	0.28	0.872	-0.21	0.51	0.139	23
Cont.Females	24	166.96	8.93	168.71	7.97	1.75	0.77	1.444							

	January Diff. in Means Female minus Male Diff.	t-Ratio	May Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	6.52	2.532*	6.92	3.254*	38
Controls	5.96	2.272*	6.21	2.625*	38

*Significant at 0.05 level (two-tailed test).

TABLE D-VIII(3)

THE PERFORMANCE OF 40 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 3

Subgroup	N	Cooperative Eng. Test Converted Score September Mean	S. D.	May Mean	S. D.	Difference in Means, May minus September	r	t-Ratio	Difference in Sept. Means Control minus Experimental Diff.	r	t	Difference in May Means Control minus Experimental Diff.	r	t	Degrees of Freedom
Experimental	40	162.35	7.76	166.15	7.26	3.80	0.79	4.884*	-0.63	0.79	0.742	0.07	0.45	0.060	39
Control	40	161.73	8.26	166.23	7.76	4.50	0.71	4.568*							
Exp.Males	16	159.13	7.65	162.00	6.64	2.87	0.68	1.914	-2.56	0.68	1.657	0.50	0.00	0.221	15
Cont.Males	16	156.56	7.22	162.50	5.69	5.94	0.39	3.176*							
Exp.Females	24	164.50	7.06	168.92	6.26	4.42	0.83	5.324*	0.67	0.82	0.754	-0.21	0.51	0.139	23
Cont.Females	24	165.17	7.02	168.71	7.97	3.54	0.78	3.342*							

	September Diff. in Means Female minus Male Diff.	t-Ratio	May Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	5.37	2.224*	6.92	3.254*	38
Controls	8.61	3.661*	6.21	2.625*	38

*Significant at 0.05 level (two-tailed test).

TABLE D-IX(3)

THE PERFORMANCE OF 40 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 3

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in September Means, Control minus Experimental			Diff. in January Means, Control minus Experimental			
		September		January		Jan. minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	40	469.00	78.03	494.48	82.75	25.47	0.70	2.532*	5.95	0.70	0.640	-9.82	0.52	0.809	39
Control	40	474.95	69.86	484.65	71.10	9.70	0.64	1.011							
Exp.Males	16	431.63	61.04	442.63	57.21	11.00	0.52	0.734	25.44	0.31	1.423	21.12	0.21	1.218	15
Cont.Males	16	457.06	57.07	463.75	48.95	6.69	0.62	0.551							
Exp.Females	24	493.92	78.18	529.04	79.01	35.13	0.67	2.630*	-7.04	0.83	0.755	-30.46	0.56	1.968	23
Cont.Females	24	486.88	74.89	498.58	79.64	11.71	0.62	0.837							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	62.29	2.619*	86.41	3.671*	38
Controls	29.82	1.318	34.83	1.524	38

*Significant at 0.05 level (two-tailed test).

TABLE D-X(3)

THE PERFORMANCE OF 40 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 3

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in January Means, Control minus Experimental			Diff. in May Means, Control minus Experimental			
		January		May		May minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	January	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	40	494.48	82.75	497.85	91.03	3.37	0.66	0.292	-9.82	0.52	0.809	-2.30	0.41	0.154	39
Control	40	484.65	71.10	495.55	79.69	10.90	0.55	0.945							
Exp.Males	16	442.63	57.21	473.50	73.45	30.87	0.55	1.884	21.12	0.21	1.218	2.81	0.38	0.137	15
Cont.Males	16	463.75	48.95	476.31	69.66	12.56	0.25	0.655							
Exp.Females	24	529.04	79.01	514.08	97.76	-14.96	0.69	0.998	-30.46	0.56	1.968	-5.71	0.38	0.271	23
Cont.Females	24	498.58	79.64	508.38	83.29	9.79	0.63	0.666							

	January Diff. in Means Female minus Male		May Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	86.41	3.671*	40.58	1.380	38
Controls	34.83	1.524	32.07	1.239	38

*Significant at 0.05 level (two-tailed test).

TABLE D-XI(3)

THE PERFORMANCE OF 40 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 3

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in September Means, Control minus Experimental			Diff. in May Means, Control minus Experimental			
		September		May		May minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	40	469.00	78.03	497.85	91.03	28.85	0.61	2.372*	5.95	0.70	0.640	-2.30	0.41	0.154	39
Control	40	474.95	69.86	495.55	79.69	20.60	0.67	2.087*							
Exp.Males	16	431.63	61.04	473.50	73.45	41.87	0.49	2.355*	25.44	0.31	1.423	2.81	0.38	0.137	15
Cont.Males	16	457.06	57.07	476.31	69.66	19.25	0.41	1.074							
Exp.Females	24	493.92	78.18	514.08	97.76	20.17	0.62	1.222	-7.04	0.83	0.755	-5.71	0.38	0.271	23
Cont.Females	24	486.88	74.89	508.38	83.29	21.50	0.76	1.846							

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	62.29	2.619*	40.58	1.380	38
Controls	29.82	1.318	32.07	1.239	38

*Significant at 0.05 level (two-tailed test).

TABLE D-XII(3)

THE PERFORMANCE OF 40 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 3

Subgroup	N	Degrees of Freedom	September Theme Total Mean	September Theme Total S. D.	Difference in September Means Control minus Experimental Diff.	r	t-Ratio	January Theme Total Mean	January Theme Total S. D.	Difference in January Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	40	39	8.90	2.21	0.00	1.00	0.000	8.73	1.90	-0.10	0.18	0.224
Controls	40		8.90	2.21				8.63	2.41			
Exp.Males	16	15	8.31	2.36	0.00	1.00	0.000	8.00	1.77	-0.38	0.05	0.545
Cont.Males	16		8.31	2.36				7.63	2.09			
Exp.Females	24	23	9.29	2.01	0.00	1.00	0.000	9.21	1.83	0.08	0.10	0.141
Cont.Females	24		9.29	2.01				9.29	2.37			

	September Diff. in Means Female minus Male Diff.	t-Ratio	January Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.97	1.370	1.21	2.024*	38
Controls	0.97	1.370	1.66	2.225*	38

*Significant at 0.05 level (two-tailed test).

TABLE D-XIII(3)

THE PERFORMANCE OF 40 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 3

Subgroup	N	Degrees of Freedom	January Theme Mean	January Total S. D.	Difference in January Means Control minus Experimental Diff.	r	t-Ratio	May Theme Mean	May Total S. D.	Difference in May Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	40	39	8.73	1.90	-0.10	0.18	0.224	8.40	2.47	0.48	0.33	1.106
Controls	40		8.63	2.41				8.88	2.14			
Exp.Males	16	15	8.00	1.77	-0.38	0.05	0.545	7.81	2.19	0.63	0.23	0.942
Cont.Males	16		7.63	2.09				8.44	1.94			
Exp.Females	24	23	9.21	1.83	0.08	0.10	0.141	8.79	2.57	0.38	0.34	0.654
Cont.Females	24		9.29	2.37				9.17	2.21			

	January Diff. in Means Female minus Male Diff.	t-Ratio	May Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	1.21	2.024*	0.98	1.222	38
Controls	1.66	2.225*	0.73	1.046	38

*Significant at 0.05 level (two-tailed test).

TABLE D-XIV(3)

THE PERFORMANCE OF 40 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 3

Subgroup	N	Degrees of Freedom	September Theme Total Mean	September Theme Total S. D.	Difference in September Means Control minus Experimental Diff.	r	t-Ratio	May Theme Total Mean	May Theme Total S. D.	Difference in May Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	40	39	8.90	2.21	0.00	1.00	0.000	8.40	2.47	0.48	0.33	1.106
Controls	40		8.90	2.21				8.88	2.14			
Exp.Males	16	15	8.31	2.36	0.00	1.00	0.000	7.81	2.19	0.63	0.23	0.942
Cont.Males	16		8.31	2.36				8.44	1.94			
Exp.Females	24	23	9.29	2.01	0.00	1.00	0.000	8.79	2.57	0.38	0.34	0.654
Cont.Females	24		9.29	2.01				9.17	2.21			

	September Diff. in Means Female minus Male Diff.	September Diff. in Means Female minus Male t-Ratio	May Diff. in Means Female minus Male Diff.	May Diff. in Means Female minus Male t-Ratio	Degrees of Freedom
Experimentals	0.98	1.370	0.98	1.222	38
Controls	0.98	1.370	0.73	1.046	38

TABLE D-XV(3)

THE PERFORMANCE OF 11 MATCHED PAIRS OF STUDENTS ON COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964, JANUARY 1965, MAY 1965, AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, AND t-RATIOS: UNIVERSITY 3

	Cooperative English Test Converted Score									
Sub-group	Sept. 1964 Mean	S. D.	Jan. 1964 Mean	S. D.	Diff. in Means, Jan. minus Sept.	r	t-Ratio df=10	Diff. in Jan. Means, Cont. minus Exper.	r	t-Ratio df=10
Exper.	167.18	6.04	167.82	7.07	0.64	0.81	0.481	0.00	0.51	0.000
Control	166.73	7.11	167.82	7.20	1.09	0.66	0.583			

	Cooperative English Test Converted Score									
Sub-group	Jan. 1965 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Jan.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	167.82	7.07	170.36	7.92	2.55	0.77	1.566	-1.45	0.59	0.665
Control	167.82	7.20	168.91	7.33	1.09	0.71	0.629			

	Cooperative English Test Converted Score									
Sub-group	Sept. 1964 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	167.18	6.04	170.36	7.92	3.18	0.79	2.079	Same as above		
Control	166.73	7.11	168.91	7.33	2.18	0.78	1.426			

	Cooperative English Test Converted Score									
Sub-group	May 1965 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus May	r	t-Ratio	Diff. in May 1966 Means, Cont. minus Exper.	r	t-Ratio
Exper.	170.36	7.92	166.82	9.59	-3.55	0.21	1.010	3.73	0.30	1.179
Control	168.91	7.33	170.55	6.85	1.64	0.56	0.776			

	Cooperative English Test Converted Score									
Sub-group	Sept. 1964 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May 1966 Means, Cont. minus Exper.	r	t-Ratio
Exper.	167.18	6.04	166.82	9.59	-0.36	0.27	0.116	Same as above		
Control	166.73	7.11	170.55	6.85	3.82	0.62	1.993			

TABLE D-XVI(3)

THE PERFORMANCE OF 11 MATCHED PAIRS OF STUDENTS ON THE COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TEST IN SEPTEMBER 1964, JANUARY 1965, MAY 1965, AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, AND t-RATIOS: UNIVERSITY 3

	College Entrance Exam. Board Stan. Rating									
Sub-group	Sept. 1964 Mean	S. D.	Jan. 1965 Mean	S. D.	Diff. in Means, Jan. minus Sept.	r	t-Ratio df=10	Diff. in Jan. Means, Cont. minus Exper.	r	t-Ratio df=10
Exper.	496.91	85.62	530.36	95.08	33.45	0.65	1.399	-19.91	0.43	0.698
Control	497.91	82.82	510.45	69.10	12.55	0.66	0.623			

	College Entrance Exam. Board Stan. Rating									
Sub-group	Jan. 1965 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Jan.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	530.36	95.08	520.73	61.78	- 9.64	0.95	0.728	-13.18	0.65	0.742
Control	510.45	69.10	507.55	71.08	- 2.91	0.66	0.160			

	College Entrance Exam. Board Stan. Rating									
Sub-group	Sept. 1964 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	496.91	85.62	520.73	61.78	23.82	0.58	1.059	Same as above		
Control	497.91	82.82	507.55	71.08	9.64	0.92	0.932			

	College Entrance Exam. Board Stan. Rating									
Sub-group	May 1965 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus May	r	t-Ratio	Diff. in May 1966 Means, Cont. minus Exper.	r	t-Ratio
Exper.	520.73	61.78	567.55	94.07	46.82	0.76	2.389*	- 8.73	0.71	0.406
Control	507.55	71.08	558.82	79.13	51.27	0.73	2.918*			

	College Entrance Exam. Board Stan. Rating									
Sub-group	Sept. 1964 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May 1966 Means, Cont. minus Exper.	r	t-Ratio
Exper.	496.91	85.62	567.55	94.07	70.64	0.84	4.317*	Same as above		
Control	497.91	82.82	558.82	79.13	60.91	0.72	3.202*			

*Significant at 0.05 level (two-tailed test).

TABLE D-XVII(3)

THE PERFORMANCE OF 11 MATCHED PAIRS OF STUDENTS ON THE TOTAL OF TWO THEME RATINGS IN SEPTEMBER 1964, JANUARY 1965, MAY 1965, AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, AND t-RATIOS: UNIVERSITY 3

		Theme Rating 1		Theme Rating 2		Theme Rating Total			Diff. in Means on Theme Total			
Date	Subgroup	Mean	S. D.	Mean	S. D.	Mean	S. D.	r, Reader 1 with Reader 2	Cont.minus Exp.	r, Cont. with Exp.	t-Ratio	d.f.
9/64	Exper.	Not available				9.09	1.44	Not available	0.00	1.00	0.000	10
9/64	Cont.					9.09	1.44					
1/65	Exper.	4.55	1.16	4.64	1.23	9.18	1.99	0.40	0.82	0.08	1.000	10
1/65	Cont.	5.09	1.16	4.91	1.24	10.00	1.81	0.13				
5/65	Exper.	4.73	1.29	4.91	1.50	9.64	2.35	0.41	0.27	0.53	0.375	10
5/65	Cont.	4.73	1.21	5.18	1.59	9.91	2.39	0.45				
5/66	Exper.	3.91	1.08	4.64	0.77	8.55	1.44	0.18	-0.82	-0.04	1.014	10
5/66	Cont.	3.82	1.40	3.91	0.90	7.73	2.05	0.56				

LIST OF TABLES
University 4

Table Page

D-I(4) Achievement as of September 1964 of a Sample of 1964-65 Freshman Students and of Various Persisting Portions of that Sample: University 4 259

D-II(4) Performance of Available Samples of Matched Pairs of Students on Three Criterion Measures at Beginning and Middle of First Year of College: University 4 261

D-III(4) The Performance of 20 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 4 262

D-IV(4) The Performance of 20 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 4 263

D-V(4) The Performance of 20 Matched Pairs of Students on Total Theme Rating in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 4 264

TABLE D-I(4)

ACHIEVEMENT AS OF SEPTEMBER 1964 OF A SAMPLE OF 1964-65 FRESHMAN STUDENTS AND OF VARIOUS PERSISTING PORTIONS OF THAT SAMPLE: UNIVERSITY 4

Sample and Subgroups	N	Percent Men	Percentile Rank in H. S. Class Mean	Percentile Rank in H. S. Class S. D.	ACT English Stan. Score Mean	ACT English Stan. Score S. D.	ACT Composite Stan. Score Mean	ACT Composite Stan. Score S. D.	COOP English Exp. Converted Score Mean	COOP English Exp. Converted Score S. D.
Exp. Pool plus Control Pool	946	45.2	61.26	20.47	21.06	4.00	22.01	3.73	160.34	9.05
Exp. Pool	323	46.7	60.62	20.57	21.08	3.91	22.09	3.74	160.77	8.97
Control Pool	623	44.5	61.59	20.40	21.05	4.04	21.97	3.72	160.12	9.08
Matched Exp. September 1964	252	43.3	60.95	20.45	21.17	3.70	22.13	3.55	160.81	8.57
Matched Controls September 1964	252	43.3	62.22	19.96	21.20	3.91	22.10	3.63	160.83	8.07
Exp. Group January 1965	20	45.0	60.50	20.37	21.15	3.37	22.75	3.28	160.80	7.98
Control Group January 1965	20	45.0	56.00	20.59	21.25	3.63	21.60	3.25	160.55	8.76

TABLE D-I(4)
CONTINUED

Sample and Subgroups	N	CEEB English Composition Stan. Rating		Z-Score[1]		Theme Rating 1		September Theme Theme Rating 2		Theme Rating Total	
		Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.
Exp. Pool plus Control Pool	946	462.27	97.82	99.50	18.54	4.65	1.65	4.43	1.55	9.09	2.68
Exp. Pool	323	466.15	97.65	100.37	18.38	4.76	1.64	4.45	1.53	9.21	2.64
Control Pool	623	460.26	97.85	99.05	18.61	4.59	1.66	4.43	1.56	9.02	2.70
Matched Exp. September 1964	252	467.55	89.63	100.55	16.94	4.78	1.52	4.41	1.43	9.19	2.37
Matched Controls September 1964	252	468.70	89.80	100.70	16.72	4.67	1.52	4.51	1.42	9.19	2.37
Exp. Group January 1965	20	470.75	75.66	100.95	14.42	4.90	1.45	4.65	1.35	9.55	2.18
Control Group January 1965	20	474.30	57.83	101.05	14.46	5.15	1.65	4.40	1.16	9.55	2.18

---

[1] Combination of Cooperative English Tests: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

TABLE D-II(4)

PERFORMANCE OF AVAILABLE SAMPLES OF MATCHED PAIRS OF STUDENTS ON THREE CRITERION MEASURES AT BEGINNING AND MIDDLE OF FIRST YEAR OF COLLEGE: UNIVERSITY 4

Subgroup	N	Time of Testing	Cooperative English Tests: English Expression (1960) Converted Score		CEEB English Composition Test Standard Rating		Theme Rating 1		Theme Rating 2		Theme Rating Total	
			Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.
Experimental	252	Sept.1964	160.81	8.57	467.55	89.63	4.78	1.52	4.41	1.43	9.19	2.37
Control	252	"	160.83	8.07	468.70	89.80	4.67	1.52	4.51	1.42	9.19	2.37
Experimental	20	Jan. 1965	156.05	14.17	497.70	66.61	5.00	1.52	4.80	1.36	9.80	2.64
Control	20	"	161.55	11.49	495.85	80.16	5.05	1.28	5.00	1.87	10.05	2.42

TABLE D-III(4)

THE PERFORMANCE OF 20 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 4

		Cooperative Eng. Test Converted Score				Difference in Means,			Difference in Sept. Means Control minus Experimental			Difference in Jan. Means Control minus Experimental			
		September		January		Jan. minus									Degrees of
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t	Diff.	r	t	Freedom
Experimental	20	160.80	7.98	156.05	14.17	-4.75	0.29	1.469	-0.25	0.83	0.220	5.50	-0.17	1.216	19
Control	20	160.55	8.76	161.55	11.49	1.00	0.44	0.397							
Exp.Males	9	158.78	8.23	145.56	11.82	-13.22	-0.23	2.355*	-1.44	0.94	1.100	15.56	-0.36	2.852*	8
Cont.Males	9	157.33	10.03	161.11	6.52	3.78	0.80	1.710							
Exp.Females	11	162.45	7.38	164.64	9.31	2.18	0.68	0.994	0.73	0.68	0.411	-2.73	-0.26	0.453	10
Cont.Females	11	163.18	6.46	161.91	14.32	-1.27	0.39	0.304							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		Degrees of
	Diff.	t-Ratio	Diff.	t-Ratio	Freedom
Experimentals	3.67	0.998	19.08	3.831*	18
Controls	5.85	1.494	0.80	0.147	18

*Significant at 0.05 level (two-tailed test).

TABLE D-IV(4)

THE PERFORMANCE OF 20 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 4

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in September Means, Control minus Experimental			Diff. in January Means, Control minus Experimental			
		September		January		Jan. minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	20	470.75	75.66	497.70	66.61	26.95	0.47	1.598	3.55	0.74	0.305	-1.85	0.32	0.094	19
Control	20	474.30	57.83	495.85	80.16	21.55	0.39	1.193							
Exp.Males	9	448.11	89.26	479.22	72.47	31.11	0.51	1.087	13.33	0.93	0.932	11.78	0.47	0.484	8
Cont.Males	9	461.44	58.88	491.00	59.39	29.56	0.18	1.103							
Exp.Females	11	489.27	55.92	512.82	57.11	23.55	0.31	1.118	-4.45	0.48	0.249	-13.00	0.25	0.425	10
Cont.Females	11	484.82	54.75	499.82	93.61	15.00	0.52	0.589							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	41.16	1.193	33.60	1.100	18
Controls	23.38	0.871	8.82	0.233	18

TABLE D-V(4)

THE PERFORMANCE OF 20 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 4

			September Theme Total		Difference in September Means Control minus Experimental			January Theme Total		Difference in January Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Mean	S. D.	Diff.	r	t-Ratio	Mean	S. D.	Diff.	r	t-Ratio
Experimentals	20	19	9.55	2.18	0.00	1.00	0.000	9.80	2.64	0.25	0.51	0.435
Controls	20		9.55	2.18				10.05	2.42			
Exp.Males	9	8	8.78	2.10	0.00	1.00	0.000	8.22	2.90	1.33	0.57	1.569
Cont.Males	9		8.78	2.10				9.56	1.95			
Exp.Females	11	10	10.18	2.04	0.00	1.00	0.000	11.09	1.44	-0.64	0.55	0.904
Cont.Females	11		10.18	2.04				10.45	2.68			

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	1.40	1.436	2.87	2.729*	18
Controls	1.40	1.436	0.89	0.798	18

*Significant at 0.05 level (two-tailed test).

LIST OF TABLES
University 5

Table		Page
D-I(5)	Achievement as of September 1964 of a Sample of 1964-65 Freshman Students and of Various Persisting Portions of that Sample: University 5	269
D-II(5)	Performance of Available Samples of Matched Pairs of Students on Three Criterion Measures at Beginning, Middle, and End of First Year and End of Second Year of College: University 5	271
D-III(5)	The Performance of 126 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 5	272
D-IV(5)	The Performance of 126 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 5	273
D-V(5)	The Performance of 126 Matched Pairs of Students on Total Theme Rating in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 5	274
D-VI(5)	The Performance of 90 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 5	275
D-VII(5)	The Performance of 90 Matched Pairs of Students on the Cooperative English Tests: English Expression in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 5	276

Table		Page
D-VIII(5)	The Performance of 90 Matched Pairs of Students on the Cooperative English Tests: English Expression in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 5	277
D-IX(5)	The Performance of 90 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 5	278
D-X(5)	The Performance of 90 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 5	279
D-XI(5)	The Performance of 90 Matched Pairs of Students on College Entrance Examination Board English Composition Tests in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 5	280
D-XII(5)	The Performance of 90 Matched Pairs of Students on Total Theme Rating in September 1964 and January 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 5	281
D-XIII(5)	The Performance of 90 Matched Pairs of Students on Total Theme Rating in January 1965 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 5	282
D-XIV(5)	The Performance of 90 Matched Pairs of Students on Total Theme Rating in September 1964 and May 1965; Including Differences in Means, Standard Deviations, t-Ratios, and Comparisons Between the Sexes: University 5	283

Table Page

D-XV(5) The Performance of 45 Matched Pairs of Students on Cooperative English Tests: English Expression in September 1964, January 1965, May 1965, and May 1966; Including Differences in Means, Standard Deviations, and t-Ratios: University 5 284

D-XVI(5) The Performance of 45 Matched Pairs of Students on the College Entrance Examination Board English Composition Test in September 1964, January 1965, May 1965, and May 1966; Including Differences in Means, Standard Deviations, and t-Ratios: University 5 285

D-XVII(5) The Performance of 45 Matched Pairs of Students on the Total of Two Theme Ratings in September 1964, January 1965, May 1965, and May 1966; Including Differences in Means, Standard Deviations, and t-Ratios: University 5 286

TABLE D-I(5)

ACHIEVEMENT AS OF SEPTEMBER 1964 OF A SAMPLE OF 1964-65 FRESHMAN STUDENTS AND OF VARIOUS PERSISTING PORTIONS OF THAT SAMPLE: UNIVERSITY 5

Sample and Subgroups	N	Percent Men	Percentile Rank in H. S. Class		ACT English Stan. Score		ACT Composite Stan. Score		COOP English Exp. Converted Score	
			Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.
Exp. Pool plus Control Pool	705	49.2	67.96	21.09	22.85	3.40	24.68	3.29	162.16	8.86
Exp. Pool	227	50.4	68.77	21.18	22.93	3.41	24.89	3.44	162.70	8.53
Control Pool	477	50.9	67.61	21.05	22.82	3.39	24.59	3.22	161.90	9.01
Matched Exp. September 1964	162	51.2	67.96	21.81	22.73	3.18	24.72	3.25	161.93	6.83
Matched Controls September 1964	162	51.2	66.98	21.26	22.67	3.06	24.56	3.01	162.62	6.81
Exp. Group January 1965	126	49.2	70.32	19.60	23.09	3.16	25.13	3.17	162.12	6.94
Control Group January 1965	126	49.2	67.30	21.17	22.82	3.12	24.60	3.05	162.91	6.84
Exp. Group May 1965	90	46.7	70.11	20.25	23.32	3.22	25.31	3.22	161.98	7.33
Control Group May 1965	90	46.7	66.89	21.33	22.86	3.09	24.30	3.01	163.17	6.98
Exp. Group May 1966	45	51.1	70.00	23.38	Not Available				161.56	8.00
Control Group May 1966	45	51.1	68.44	19.77	"				163.49	7.42

TABLE D-I(5)
CONTINUED

		CEEB English Composition Stan. Rating		Z-Score[1]		Theme Rating 1		September Theme Theme Rating 2		Theme Rating Total	
Sample and Subgroups	N	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.
Exp. Pool plus Control Pool	705	477.28	99.24	99.51	18.45	4.57	1.44	4.32	1.47	8.89	2.42
Exp. Pool	227	485.76	99.22	100.83	18.26	4.64	1.45	4.41	1.54	9.04	2.44
Control Pool	477	473.39	99.05	98.91	18.53	4.54	1.43	4.28	1.43	8.83	2.41
Matched Exp. September 1964	162	482.45	81.75	99.48	14.01	4.58	1.24	4.35	1.31	8.93	1.88
Matched Controls September 1964	162	475.43	80.22	99.57	13.85	4.57	1.18	4.36	1.20	8.93	1.88
Exp. Group January 1965	126	486.59	83.15	100.11	14.41	4.61	1.26	4.37	1.34	8.98	1.84
Control Group January 1965	126	478.03	83.01	100.19	14.19	4.62	1.20	4.36	1.14	8.98	1.84
Exp. Group May 1965	90	486.99	80.93	99.98	14.78	4.57	1.27	4.48	1.33	9.04	1.78
Control Group May 1965	90	475.12	84.17	100.17	14.54	4.64	1.22	4.39	1.09	9.04	1.78
Exp. Group May 1966	45	493.76	85,57	100.18	15.75	Not Available				9.13	1.75
Control Group May 1966	45	475.91	86.00	100.58	15.44	"				9.13	1.75

---

[1] Combination of Cooperative English Tests: English Expression and College Entrance Examination Board English Composition Test, September scores. See page 39.

TABLE D-II(5)

PERFORMANCE OF AVAILABLE SAMPLES OF MATCHED PAIRS OF STUDENTS ON THREE CRITERION MEASURES AT BEGINNING, MIDDLE, AND END OF FIRST YEAR AND END OF SECOND YEAR OF COLLEGE: UNIVERSITY 5

			Cooperative English Tests: English Expression (1960) Converted Score		CEEB English Composition Test Standard Rating		Theme Rating 1		Theme Rating 2		Theme Rating Total	
Subgroup	N	Time of Testing	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S. D.	Mean	S.D.
Experimental	162	Sept.1964	161.93	6.83	482.45	81.75	4.58	1.24	4.35	1.31	8.93	1.88
Control	162	"	162.62	6.81	475.43	80.22	4.57	1.18	4.36	1.20	8.93	1.88
Experimental	126	Jan. 1965	163.58	6.95	522.66	87.96	4.85	1.51	5.06	1.49	9.91	2.38
Control	126	"	163.89	6.63	511.94	95.62	4.83	1.36	5.19	1.47	10.02	2.29
Experimental	90	May 1965	167.22	6.32	526.06	79.31	4.98	1.30	4.83	1.34	9.81	2.05
Control	90	"	167.88	6.67	517.91	72.45	4.92	1.34	5.03	1.31	9.96	2.04
Experimental	45	May 1966	164.80	9.60	559.40	80.01	4.80	1.42	4.78	1.47	9.58	2.58
Control	45	"	166.64	7.76	545.75	79.41	4.82	1.32	4.98	1.51	9.80	2.55

TABLE D-III(5)

THE PERFORMANCE OF 126 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 5

Subgroup	N	Cooperative Eng. Test Converted Score September Mean	S. D.	January Mean	S. D.	Difference in Means, Jan. minus September	r	t-Ratio	Difference in Sept.Means Control minus Experimental Diff.	r	t	Difference in Jan. Means Control minus Experimental Diff.	r	t	Degrees of Freedom
Experimental	126	162.12	6.94	163.58	6.95	1.46	0.65	2.828*	0.79	0.75	1.815	0.31	0.54	0.531	125
Control	126	162.91	6.84	163.89	6.63	0.98	0.67	1.988*							
Exp.Males	62	160.82	6.01	161.92	6.45	1.10	0.51	1.390	0.87	0.64	1.229	0.27	0.47	0.308	61
Cont.Males	62	161.69	6.93	162.19	7.03	0.50	0.61	0.632							
Exp.Females	64	163.38	7.53	165.19	7.04	1.81	0.73	2.692*	0.72	0.83	1.368	0.34	0.57	0.450	63
Cont.Females	64	164.09	6.53	165.53	5.76	1.44	0.72	2.436*							

	September Diff. in Means Female minus Male Diff.	t-Ratio	January Diff.in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	2.56	2.081*	3.27	2.692*	124
Controls	2.40	1.984*	3.34	2.894*	124

*Significant at 0.05 level (two-tailed test).

TABLE D-IV(5)

THE PERFORMANCE OF 126 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 5

Subgroup	N	College Entrance Examination Board Standard Rating September Mean	S. D.	January Mean	S. D.	Difference in Means Jan. minus September	r	t-Ratio	Diff. in September Means, Control minus Experimental Diff.	r	t-Ratio	Diff. in January Means, Control minus Experimental Diff.	r	t-Ratio	d.f.
Experimental	126	486.59	83.15	522.66	87.96	36.07	0.68	5.887*	-8.56	0.79	1.760	-10.71	0.42	1.272	125
Control	126	478.03	83.01	511.94	95.62	33.91	0.67	5.178*							
Exp.Males	62	465.42	87.52	508.10	89.78	42.68	0.65	4.515*	-9.95	0.71	1.252	-23.56	0.34	1.844	61
Cont.Males	62	455.47	72.38	484.53	83.25	29.06	0.50	2.896*							
Exp.Females	64	507.09	73.05	536.77	83.79	29.67	0.69	3.783*	-7.20	0.85	1.255	1.73	0.56	0.159	63
Cont.Females	64	499.89	86.71	538.50	99.27	38.61	0.74	4.538*							

	September Diff. in Means Female minus Male Diff.	t-Ratio	January Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	41.67	2.882*	28.67	1.839	124
Controls	44.42	3.092*	53.97	3.275*	124

*Significant at 0.05 level (two-tailed test).

TABLE D-V(5)

THE PERFORMANCE OF 126 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 5

			September Theme Total		Difference in September Means Control minus Experimental			January Theme Total		Difference in January Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Mean	S. D.	Diff.	r	t-Ratio	Mean	S. D.	Diff.	r	t-Ratio
Experimentals	126	125	8.98	1.84	0.00	1.00	0.000	9.91	2.38	0.10	0.38	0.443
Controls	126		8.98	1.84				10.02	2.28			
Exp.Males	62	61	8.61	1.90	0.00	1.00	0.000	9.13	2.52	0.45	0.29	1.231
Cont.Males	62		8.61	1.90				9.58	2.29			
Exp.Females	64	63	9.34	1.70	0.00	1.00	0.000	10.67	1.94	-0.23	0.41	0.820
Cont.Females	64		9.34	1.70				10.44	2.20			

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.73	2.258*	1.54	3.820*	124
Controls	0.73	2.258*	0.86	2.125*	124

*Significant at 0.05 level (two-tailed test).

TABLE D-VI(5)

THE PERFORMANCE OF 90 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 5

Subgroup	N	Cooperative Eng. Test Converted Score September Mean	September S. D.	January Mean	January S. D.	Difference in Means, Jan. minus September	r	t-Ratio	Difference in Sept. Means Control minus Experimental Diff.	r	t	Difference in Jan. Means Control minus Experimental Diff.	r	t	Degrees of Freedom
Experimental	90	161.98	7.33	163.50	7.38	1.52	0.66	2.372*	1.19	0.78	2.365*	0.34	0.59	0.515	89
Control	90	163.17	6.98	163.84	6.42	0.68	0.61	1.070							
Exp.Males	42	160.52	6.69	162.12	7.32	1.60	0.53	1.501	1.62	0.72	2.004	-0.48	0.59	0.481	41
Cont.Males	42	162.14	7.15	161.64	6.59	-0.50	0.50	0.464							
Exp.Females	48	163.25	7.63	164.71	7.21	1.46	0.75	1.884	0.81	0.83	1.297	1.06	0.56	1.176	47
Cont.Females	48	164.06	6.71	165.77	5.59	1.71	0.71	2.444*							

	September Diff. in Means Female minus Male Diff.	t-Ratio	January Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	2.73	1.770	2.59	1.668	88
Controls	1.92	1.299	4.13	3.179*	88

*Significant at 0.05 level (two-tailed test).

TABLE D-VII(5)

THE PERFORMANCE OF 90 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 5

Subgroup	N	Cooperative Eng. Test Converted Score January Mean	January S. D.	May Mean	May S. D.	Difference in Means, May minus January	r	t-Ratio	Difference in Jan. Means Control minus Experimental Diff.	r	t	Difference in May Means Control minus Experimental Diff.	r	t	Degrees of Freedom
Experimental	90	163.50	7.38	167.22	6.32	3.72	0.69	6.411*	0.34	0.59	0.515	0.66	0.69	1.208	89
Control	90	163.84	6.42	167.88	6.67	4.03	0.74	8.002*							
Exp.Males	42	162.12	7.32	166.60	5.45	4.48	0.75	5.908*	-0.48	0.59	0.481	-0.50	0.57	0.624	41
Cont.Males	42	161.64	6.59	166.10	5.65	4.45	0.68	5.728*							
Exp.Females	48	164.71	7.21	167.77	6.94	3.06	0.65	3.563*	1.06	0.56	1.176	1.67	0.76	2.338*	47
Cont.Females	48	165.77	5.59	169.44	7.09	3.67	0.77	5.561*							

	January Diff. in Means Female minus Male Diff.	t-Ratio	May Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	2.59	1.668	1.17	0.874	88
Controls	4.13	3.179*	3.34	2.421*	88

*Significant at 0.05 level (two-tailed test).

TABLE D-VIII(5)

THE PERFORMANCE OF 90 MATCHED PAIRS OF STUDENTS ON THE COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 5

		Cooperative Eng. Test Converted Score September		May		Difference in Means, May minus September			Difference in Sept. Means Control minus Experimental			Difference in May Means Control minus Experimental			Degrees of Freedom
Subgroup	N	Mean	S. D.	Mean	S. D.		r	t-Ratio	Diff.	r	t	Diff.	r	t	
Experimental	90	161.98	7.33	167.22	6.32	5.24	0.69	9.086*	1.19	0.78	2.365*	0.66	0.69	1.208	89
Control	90	163.17	6.98	167.38	6.67	4.71	0.71	8.528*							
Exp.Males	42	160.52	6.69	166.60	5.45	6.07	0.64	7.421*	1.62	0.72	2.004	-0.50	0.57	0.624	41
Cont.Males	42	162.14	7.15	166.10	5.65	3.95	0.59	4.261*							
Exp.Females	48	163.25	7.63	167.77	6.94	4.52	0.72	5.616*	0.81	0.83	1.297	1.67	0.76	2.338*	47
Cont.Females	48	164.06	6.71	169.44	7.09	5.37	0.80	8.430*							

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	2.73	1.770	1.17	0.874	88
Controls	1.92	1.299	3.34	2.421*	88

*Significant at 0.05 level (two-tailed test).

TABLE D-IX(5)

THE PERFORMANCE OF 90 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 5

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in September Means, Control minus Experimental			Diff. in January Means, Control minus Experimental			
		September		January		Jan. minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	90	486.99	80.93	524.70	89.72	37.71	0.75	5.858*	-11.87	0.79	2.098*	-15.89	0.52	1.632	89
Control	90	475.12	84.17	508.81	97.44	33.69	0.65	4.130*							
Exp.Males	42	471.31	89.40	509.95	91.08	38.64	0.78	4.133*	-16.50	0.76	1.780	-33.83	0.41	2.241*	41
Cont.Males	42	454.81	79.84	476.12	87.47	21.31	0.47	1.580							
Exp.Females	48	500.71	69.89	537.60	86.47	36.90	0.71	4.115*	-7.81	0.83	1.137	-0.19	0.58	0.015	47
Cont.Females	48	492.90	83.85	537.42	96.71	44.52	0.74	4.641*							

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		Degrees of
	Diff.	t-Ratio	Diff.	t-Ratio	Freedom
Experimentals	29.40	1.729	27.65	1.460	88
Controls	38.09	2.174*	61.30	3.101*	88

*Significant at 0.05 level (two-tailed test).

TABLE D-X(5)

THE PERFORMANCE OF 90 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 5

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in January Means, Control minus Experimental			Diff. in May Means, Control minus Experimental			
		January		May		May minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	January	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	90	524.70	89.72	526.06	79.31	1.36	0.67	0.184	-15.89	0.52	1.632	-8.14	0.53	1.038	89
Control	90	508.81	97.44	517.91	72.45	9.10	0.72	1.261							
Exp.Males	42	509.95	91.08	508.45	77.12	-1.50	0.72	0.150	-33.83	0.41	2.241*	-11.14	0.50	1.003	41
Cont.Males	42	476.12	87.47	497.31	64.01	21.19	0.68	2.097*							
Exp.Females	48	537.60	86.47	541.46	77.99	3.85	0.60	0.359	-0.19	0.58	0.015	-5.52	0.50	0.497	47
Cont.Females	48	537.42	96.71	535.94	74.56	-1.48	0.70	0.147							

	January Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	27.65	1.460	33.01	1.991*	88
Controls	61.30	3.101*	38.63	2.589*	88

*Significant at 0.05 level (two-tailed test).

TABLE D-XI(5)

THE PERFORMANCE OF 90 MATCHED PAIRS OF STUDENTS ON COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TESTS IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 5

		College Entrance Examination Board Standard Rating				Difference in Means			Diff. in September Means, Control minus Experimental			Diff. in May Means, Control minus Experimental			
		September		May		May minus									
Subgroup	N	Mean	S. D.	Mean	S. D.	September	r	t-Ratio	Diff.	r	t-Ratio	Diff.	r	t-Ratio	d.f.
Experimental	90	486.99	80.93	526.06	79.31	39.07	0.66	5.568*	-11.87	0.79	2.098*	-8.14	0.53	1.038	89
Control	90	475.12	84.17	517.91	72.45	42.79	0.71	6.663*							
Exp.Males	42	471.31	89.40	508.45	77.12	37.14	0.75	3.945*	-16.50	0.76	1.780	-11.14	0.50	1.003	41
Cont.Males	42	454.81	79.84	497.31	64.01	42.50	0.64	4.361*							
Exp.Females	48	500.71	69.89	541.46	77.99	40.75	0.54	3.940*	-7.81	0.83	1.137	-5.52	0.50	0.497	47
Cont.Females	48	492.90	83.85	535.94	74.56	43.04	0.73	5.006*							

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		Degrees of Freedom
	Diff.	t-Ratio	Diff.	t-Ratio	
Experimentals	29.40	1.729	33.01	1.991*	88
Controls	38.09	2.174*	38.63	2.589*	88

*Significant at 0.05 level (two-tailed test).

TABLE D-XII(5)

THE PERFORMANCE OF 90 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND JANUARY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 5

			September Theme Total		Difference in September Means Control minus Experimental			January Theme Total		Difference in January Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Mean	S. D.	Diff.	r	t-Ratio	Mean	S. D.	Diff.	r	t-Ratio
Experimentals	90	89	9.04	1.78	0.00	1.00	0.000	9.80	2.31	0.20	0.35	0.705
Controls	90		9.04	1.78				10.00	2.37			
Exp.Males	42	41	8.83	1.86	0.00	1.00	0.000	8.98	2.39	0.62	0.24	1.362
Cont.Males	42		8.83	1.86				9.60	2.34			
Exp.Females	48	47	9.23	1.69	0.00	1.00	0.000	10.52	1.98	-0.17	0.40	0.478
Cont.Females	48		9.23	1.69				10.35	2.35			

5

	September Diff. in Means Female minus Male		January Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.40	1.046	1.54	3.311*	88
Controls	0.40	1.046	0.75	1.514	88

*Significant at 0.05 level (two-tailed test).

TABLE D-XIII(5)

THE PERFORMANCE OF 90 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN JANUARY 1965 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 5

Subgroup	N	Degrees of Freedom	January Theme Mean	January Total S. D.	Difference in January Means Control minus Experimental Diff.	r	t-Ratio	May Theme Mean	May Total S. D.	Difference in May Means Control minus Experimental Diff.	r	t-Ratio
Experimentals	90	89	9.80	2.31	0.20	0.35	0.705	9.81	2.05	0.14	0.17	0.516
Controls	90		10.00	2.37				9.96	2.04			
Exp.Males	42	41	8.98	2.39	0.62	0.24	1.362	9.38	1.81	0.26	0.28	0.696
Cont.Males	42		9.60	2.34				9.64	2.17			
Exp.Females	48	47	10.52	1.98	-0.17	0.40	0.478	10.19	2.18	0.04	0.04	0.101
Cont.Females	48		10.35	2.35				10.23	1.87			

	January Diff. in Means Female minus Male Diff.	t-Ratio	May Diff. in Means Female minus Male Diff.	t-Ratio	Degrees of Freedom
Experimentals	1.54	3.311*	0.81	1.874	88
Controls	0.75	1.514	0.59	1.360	88

*Significant at 0.05 level (two-tailed test).

TABLE D-XIV(5)

THE PERFORMANCE OF 90 MATCHED PAIRS OF STUDENTS ON TOTAL THEME RATING IN SEPTEMBER 1964 AND MAY 1965; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, t-RATIOS, AND COMPARISONS BETWEEN THE SEXES: UNIVERSITY 5

			September		Difference in September Means Control minus Experimental			May		Difference in May Means Control minus Experimental		
Subgroup	N	Degrees of Freedom	Theme Mean	Total S. D.	Diff.	r	t-Ratio	Theme Mean	Total S. D.	Diff.	r	t-Ratio
Experimentals	90	89	9.04	1.78	0.00	1.00	0.000	9.81	2.05	0.14	0.17	0.516
Controls	90		9.04	1.78				9.96	2.04			
Exp.Males	42	41	8.83	1.86	0.00	1.00	0.000	9.38	1.81	0.26	0.28	0.696
Cont.Males	42		8.83	1.86				9.64	2.17			
Exp.Females	48	47	9.23	1.69	0.00	1.00	0.000	10.19	2.18	0.04	0.04	0.101
Cont.Females	48		9.23	1.69				10.23	1.87			

	September Diff. in Means Female minus Male		May Diff. in Means Female minus Male		
	Diff.	t-Ratio	Diff.	t-Ratio	Degrees of Freedom
Experimentals	0.40	1.046	0.81	1.874	88
Controls	0.40	1.046	0.59	1.360	88

*Significant at 0.05 level (two-tailed test).

TABLE D-XV(5)

THE PERFORMANCE OF 45 MATCHED PAIRS OF STUDENTS ON COOPERATIVE ENGLISH TESTS: ENGLISH EXPRESSION IN SEPTEMBER 1964, JANUARY 1965, MAY 1965, AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, AND t-RATIOS: UNIVERSITY 5

	Cooperative English Test Converted Score									
Sub-group	Sept. 1964 Mean	S. D.	Jan. 1965 Mean	S. D.	Diff. in Means, Jan. minus Sept.	r	t-Ratio df=44	Diff. in Jan. Means, Cont. minus Exper.	r	t-Ratio df=44
Exper.	161.56	8.00	163.69	6.88	2.13	0.62	2.149*	0.31	0.56	0.324
Control	163.49	7.42	164.00	6.65	0.51	0.65	0.569			

	Cooperative English Test Converted Score									
Sub-group	Jan. 1965 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Jan.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	163.69	6.88	167.73	6.43	4.04	0.65	4.789*	0.96	0.71	1.268
Control	164.00	6.65	168.69	6.62	4.69	0.74	6.477*			

	Cooperative English Test Converted Score									
Sub-group	Sept. 1964 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	161.56	8.00	167.73	6.43	6.18	0.66	6.742*	Same as above		
Control	163.49	7.42	168.69	6.62	5.20	0.67	6.004*			

	Cooperative English Test Converted Score									
Sub-group	May 1965 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus May	r	t-Ratio	Diff. in May 1966 Means, Cont. minus Exper.	r	t-Ratio
Exper.	167.73	6.43	164.80	9.60	-2.93	0.68	2.773*	1.84	0.38	1.253
Control	168.69	6.62	166.64	7.76	-2.04	0.75	2.592*			

	Cooperative English Test Converted Score									
Sub-group	Sept. 1964 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May 1966 Means, Cont. minus Exper.	r	t-Ratio
Exper.	161.56	8.00	164.80	9.60	3.24	0.63	2.810*	Same as above		
Control	163.49	7.42	166.64	7.76	3.16	0.72	3.650*			

*Significant at 0.05 level (two-tailed test).

TABLE D-XVI(5)

THE PERFORMANCE OF 45 MATCHED PAIRS OF STUDENTS ON THE COLLEGE ENTRANCE EXAMINATION BOARD ENGLISH COMPOSITION TEST IN SEPTEMBER 1964, JANUARY 1965, MAY 1965, AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, AND t-RATIOS: UNIVERSITY 5

Sub-group	College Entrance Exam. Board Stan. Rating Sept. 1964 Mean	S. D.	Jan. 1965 Mean	S. D.	Diff. in Means, Jan. minus Sept.	r	t-Ratio df=44	Diff. in Jan. Means, Cont. minus Exper.	r	t-Ratio
Exper.	493.76	85.57	534.13	83.46	40.38	0.70	4.097*	-23.58	0.55	1.758
Control	475.91	86.00	510.56	100.68	34.64	0.77	3.552*			

Sub-group	College Entrance Exam. Board Stan. Rating Jan. 1965 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Jan.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	534.13	83.46	535.13	71.51	1.00	0.59	0.094	-14.07	0.63	1.500
Control	510.56	100.68	521.07	73.91	10.51	0.77	1.088			

Sub-group	College Entrance Exam. Board Stan. Rating Sept. 1964 Mean	S. D.	May 1965 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May Means, Cont. minus Exper.	r	t-Ratio
Exper.	493.76	85.57	535.13	71.51	41.38	0.68	4.297*	Same as above		
Control	475.91	86.00	521.07	73.91	45.16	0.75	5.172*			

Sub-group	College Entrance Exam. Board Stan. Rating May 1965 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus May	r	t-Ratio	Diff. in May 1966 Means, Cont. minus Exper.	r	t-Ratio
Exper.	535.13	71.51	559.40	80.01	24.27	0.65	2.533*	-13.64	0.50	1.138
Control	521.07	73.91	545.76	79.41	24.69	0.73	2.906*			

Sub-group	College Entrance Exam. Board Stan. Rating Sept. 1964 Mean	S. D.	May 1966 Mean	S. D.	Diff. in Means, May minus Sept.	r	t-Ratio	Diff. in May 1966 Means, Cont. minus Exper.	r	t-Ratio
Exper.	493.76	85.57	559.40	80.01	65.64	0.56	5.598*	Same as above		
Control	475.91	86.00	545.76	79.41	69.84	0.65	6.689*			

*Significant at 0.05 level (two-tailed test).

TABLE D-XVII(5)

THE PERFORMANCE OF 45 MATCHED PAIRS OF STUDENTS ON THE TOTAL OF TWO THEME RATINGS IN SEPTEMBER 1964, JANUARY 1965, MAY 1965, AND MAY 1966; INCLUDING DIFFERENCES IN MEANS, STANDARD DEVIATIONS, AND t-RATIOS: UNIVERSITY 5

		Theme Rating 1		Theme Rating 2		Theme Rating Total						
Date	Subgroup	Mean	S. D.	Mean	S. D.	Mean	S. D.	r, Reader 1 with Reader 2	Diff. in Means on Theme Total Cont.minus Exp.	r, Cont. with Exp.	t-Ratio	d.f.
9/64	Exper.	Not available				9.13	1.75	Not available	0.00	1.00	0.000	44
9/64	Cont.					9.13	1.75					
1/65	Exper.	4.60	1.44	5.13	1.34	9.73	2.04	0.07	0.40	0.28	1.040	44
1/65	Cont.	4.82	1.29	5.31	1.46	10.13	2.20	0.28				
5/65	Exper.	5.07	1.31	4.80	1.38	9.87	2.05	0.17	0.22	0.26	0.591	44
5/65	Cont.	4.96	1.28	5.13	1.36	10.09	2.04	0.19				
5/66	Exper.	4.80	1.42	4.78	1.47	9.58	2.58	0.58	0.22	-0.02	0.402	44
5/66	Cont.	4.82	1.32	4.98	1.51	9.80	2.55	0.62				